UNIVERSIDADE FEDERAL DE LAVRAS

DEPARTAMENTO DE ESTATÍSTICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ESTATÍSTICA E EXPERIMENTAÇÃO AGROPECUÁRIA

JOEL JORGE NUVUNGA

MODELO FATORIAL ANALÍTICO BAYESIANO APLICADO À EXPERIMENTOS MULTI-AMBIENTES

LAVRAS, MG

2017

JOEL JORGE NUVUNGA

MODELO FATORIAL ANALÍTICO BAYESIANO APLICADO À EXPERIMENTOS MULTI-AMBIENTES

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, para obtenção do título de Doutor.

Prof. Dr. Márcio Balestre, UFLA

Orientador

Prof. Dr. Renato Ribeiro, de Lima, UFLA

Coorientador

LAVRAS-MG

2017

Ficha catalográfica elaborada pelo Sistema de Geração de Ficha Catalográfica da Biblioteca Universitária da UFLA, com dados informados pelo(a) próprio(a) autor(a).

Nuvunga, Joel Jorge.
Modelo fatorial analítico bayesiano aplicado à experimentos multi-ambiente / Joel Jorge Nuvunga. - 2017.
129 p. : il.

Orientador(a): Márcio Balestre.
Coorientador(a): Renato Ribeiro De Lima
Tese (doutorado) - Universidade Federal de Lavras, 2017.
Bibliografia.

1. Interação GEI. 2. Decomposição espectral. 3. Inferência bayesiana. I. Balestre, Márcio. II. De Lima, Renato Ribeiro. III. Título.

JOEL JORGE NUVUNGA

MODELO FATORIAL ANALÍTICO BAYESIANO APLICADO À EXPERIMENTOS MULTI-AMBIENTES

Tese apresentada à Universidade Federal de Lavras como parte das exigências do Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, área de concentração em Estatística e Experimentação Agropecuária para obtenção do título de Doutor

Aprovada em 19 de janeiro de 2017.

Dr. Júlio Sílvio de Souza Bueno Filho	UFLA
Dra. Thelma Safadi	UFLA
Dr. Renato Ribeiro de Lima	UFLA
Dr. Fernando H. R. Barrozo Toledo	CIMMYT
Dra. Alessandra Querino da Silva	UFGD

Prof. Dr. Márcio Balestre
Orientador

LAVRAS-MG

2017

Aos meus pais,
Jorge Nuvunga (In memoriam) e Tahate Cossa,
que me ensinaram a importância dos estudos
E que em todos os momentos de dificuldade,
sempre me aconselharam.

Aos meus irmãos, exemplos de perseverança,
solidariedade e pela companhia constante, amizade,
paciência e amor DEDICO.

AGRADECIMENTOS

À Universidade Federal de Lavras (UFLA) e ao Departamento de Estatística (DES), pela oportunidade concedida para a realização do Doutorado;

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), pela concessão da bolsa de Doutorado.

Ao meu orientador o Márcio Balastre, por ajudar nos meus primeiros passos no Mestrado e no Doutorado, pelo conhecimento compartilhado, confiança no meu trabalho, apoio e disposição para ajudar;

Ao Prof. Dr. Renato Ribeiro de Lima, pela amizade e apoio na orientação do trabalho;

Ao Professor Doutor Carvalho Carlos Ecole, pelo apoio incondicional para esta conquista;

Aos Profs. Drs. Júlio Sílvio de Sousa Bueno Filho, Daniel Ferreira Furtado, José Airton Rodrigues Nunes, João Domingos Scalon, Thelma Safadi, Alessandra Querino da Silva, Fernando Henrique Barroso Toledo, Marcelo Cirilo e Joel Augusto Muniz aos quais serei eternamente agradecido pela paciência, pelos ensinamentos e pela valiosa colaboração.

Aos professores do Departamento de Estatística da DES/UFLA, obrigada pela amizade e contribuição na minha formação;

Um obrigado aos meus amigos de trabalho e de luta desde o Mestrado até o momento dessa conquista que certamente sem eles seria mais difícil: Carlos Pereira da Silva ("Carlão-Programador chefe") e Luciano Antônio de Oliveira.

Aos meus colegas do doutorado, pelo constante apoio e amizade, a todos vocês que fizeram parte deste meu aprendizado e de uma forma muito especial, Jackelya Araújo da Silva, Andrezza Kellen Allan Alves, Fernando Ribeiro, Laís Mesquita, Carlos Muianga, Rafael Lemos, e Adriano Carvalho;

Aos meus irmãos; Marta, Alfredo, Matias, Jorge, Rute, Aida, Lúcia e Maria, pela amizade e companheirismo de toda vida e em especial a Rita Nuvunga e Elisar Nuvunga meu obrigado;

A todos os moçambicanos em Lavras com os quais compartilhei os melhores momentos;

A Joaquim Uate, Stelio Gadaga, Denilson Mavaieie, Carlos Balate, Júlio Miguel, João Munguambe, Rogério Romão, Herminio Bento, pela amizade e convivência, durante minha estadia em Lavras;

A turma "do bem": Lourenço Manuel, Matias Sueia Júnior, Mário Tuzine, Jonas Massuque;

A Rosélia França da Silva.

Ao Momade Álvaro, Noimilto Mindo, e a todos que colaboraram direta e indiretamente para esta conquista;

À Fundação Caloute Gulbenkian pelo apoio financeiro.

RESUMO

Um dos principais desafios presentes em programas de melhoramento de plantas é o estudo eficiente da interação entre genótipos e ambientes (GEI). A presença de interação GE significativa dificulta o trabalho do meliorista para a seleção e recomendação ampla de genótipos superiores. Dentre os diversos procedimentos estatísticos desenvolvidos para esse fim, merecem especial destaque aqueles baseados em modelos mistos via análise de fatores, comumente referidos como modelo fatorial analítico (FA). Esse consiste em uma abordagem parcimoniosa e apresenta vantagens sugestivas quando comparadas com metodologias clássicas, como a grande flexibilidade para lidar com dados desbalanceados e com variâncias heterogêneas. Entretanto, alguns problemas estão relacionados ao modelo fatorial analítico: o custo computacional em análises com grande número de ambientes e os casos *Heywood*, que torna o modelo não identificável. Além do que, a representação do modelo em *biplot* convencional não comporta nenhuma medida da incerteza referente aos escores que descrevem o efeito da GEI ou efeitos de Genótipo (G) + GEI, plotados. A proposta aqui apresentada busca descrever formas gerais de como a heterogeneidade das covariâncias genéticas e residuais podem ser modeladas na perspectiva da análise fatorial em modelos mistos, utilizando-se decomposição espectral dos efeitos genéticos dentro da ótica Bayesiana, diferente de outros procedimentos presentes na literatura, em que as cargas fatoriais são amostradas diretamente. Além disso, objetivou-se desenvolver um procedimento para incorporar inferência ao biplot, por meio da construção de regiões de credibilidade para os escores genotípicos e ambientais. Neste estudo, foram assumidas distribuições esféricas como prioris para autovetores e normais truncadas para valores singulares, além de inversas qui-quadrados escaladas para as variâncias residuais e priori não informativa para o efeito de genótipos. Essa abordagem difere dos métodos bayesianos apresentados até o momento que assumem as mesmas restrições presentes no modelo para efeitos mistos. Para exemplificar o método proposto foram usados dados simulados e dados reais cuja variável em estudo é a produtividade de espigas em $t.ha^{-1}$. As amostras para o processo de inferência foram obtidas diretamente, utilizando o amostrador de Gibbs. Realizaram-se desbalanceamentos aleatórios nos dados considerando níveis de 10%, 33% e 50% de perdas do genótipo no ambiente. De acordo com os resultados, a análise FA com duas cargas apresentou maior capacidade preditiva em relação aos modelos em competição. Desbalanceamentos de 10% e 33% apresentaram valores médios de correlação acima de 0,40 e com 50%, de 0,46. Observou-se também que o desempenho do modelo foi melhor na ordem de 50%, 33% e 10% de desbalanceamento. Verificou-se, ainda, que a análise com o modelo FA bayesiano é robusta sob grandes níveis de desbalanceamento dos dados. Um detalhe relevante nesse estudo diz respeito à seleção de modelos, que, no caso de dados reais, mostrou não ser uma tarefa trivial, necessitando de critérios adicionais. Além disso, o modelo proposto neste trabalho mostrou maior capacidade preditiva que o modelo frequentista equivalente e os parâmetros foram estimados adequadamente, sendo identificável, sem necessidade de rotacionalidade das cargas fatoriais ou de imposição de restrições, o que representa uma grande vantagem do método aqui proposto.

Palavras-chave: Interação GE. Decomposição espectral. Adaptabilidade, estabilidade, Fator Analítico. Distribuições esféricas. Inferência bayesiana.

ABSTRACT

One of the main challenges in plant breeding programs is the efficient study of the genotypes x environments interaction (GEI). The presence of significant GE interaction hinders the work of the breeder for the recommendation and selection of superior genotypes. Among the various statistical procedures developed for this purpose, special emphasis should be given to those based on mixed models through factor analysis, commonly referred to as factor-analytic model (FA). This consists of a parsimonious approach and presents suggestive advantages when compared with classical methodologies, such as the great flexibility to deal with unbalanced data and heterogeneous variances. However, some problems are related to the factor analytic model: computational cost in analyzes with large number of environments and the Heywood cases, which makes the model unidentifiable. Moreover, the model representation in conventional biplot does not include any measure of uncertainty regarding the scores that describe GEI effect or Genotype (G) + GEI effects, plotted. The present proposal seeks to describe general forms of how heterogeneity of genetic and residual covariance can be modeled from the perspective of factorial analysis in mixed models, using spectral decomposition of the genetic effects within Bayesian approach, different from other procedures present in the literature in which the factor loads are directly sampled. In addition, the objective was to develop a procedure to incorporate inference to the biplot, through the construction of regions of credibility for genotypic and environmental scores. In this study, spherical distributions were assumed as prioris for eigenvectors and truncated normal distribution for singular values, as well as scaled inverse chi-squared distribution for residual variances and non-informative priori for the effect of genotypes. This approach differs from the Bayesian methods presented so far that assume the same constraints present in the mixed effects model. To exemplify the proposed method, we used simulated data and real data which study variable is the yield of spikes in t.ha-1. Samples for the inference process were obtained directly using the Gibbs sampler. Random unbalancing was performed in the data considering levels of 10%, 33% and 50% of losses of the genotype in the environment. According to the results, the FA analysis with two loads presented higher predictive capacity than the competing models. Unbalancing of 10% and 33% had mean values of correlation above 0.40 and with 50%, of 0.46. It was also observed that the performance of the model was better in the order of 50%, 33% and 10% of imbalance. We also verified that the analysis with the Bayesian FA model is robust under large levels of data unbalance. A relevant detail in this study concerns the selection of models, which proved not to be a trivial task in the case of real data, requiring additional criteria. In addition, the model proposed in this work showed greater predictive capacity than the equivalent frequentist model and the parameters were adequately estimated, being identifiable, without the need for rotationality of factor loads or imposition of restrictions, which represents a great advantage of the method proposed here.

Keywords: GE interaction. Spectral decomposition. Adaptability, Stability, Factor Analytic. Spherical distributions. Bayesian inference.

SUMÁRIO

1. INTRODUÇÃO

A recomendação de novos genótipos de plantas para uso comercial requer confiança e predições acuradas do rendimento médio de cada cultivar. Essa informação pode ser obtida a partir da análise de uma série de ensaios de variedades em diferentes ambientes, também conhecidos como “ensaios multiambientes” (MET). Esses ensaios são necessários para isolar o efeito da interação “genótipo por ambiente” (GEI), percebida pelas respostas diferentes apresentadas pelos genótipos em função do ambiente de teste, o que geralmente dificulta o trabalho dos melhoristas na seleção e recomendação dos melhores genótipos. Surge assim, a necessidade de pesquisar métodos eficientes para identificar genótipos estáveis (que não contribuem com a interação), bem como para analisar e explorar os efeitos positivos da GEI, como adaptabilidade de genótipos a ambientes específicos, objetivando recomendações regionalizadas.

Vários modelos estatísticos têm sido propostos com essa finalidade, dentre os quais se destacam o de efeitos principais aditivos e interação multiplicativa (AMMI) e o de efeitos de genótipos mais interação (GGEbiplot) (CORNELIUS; SEYEDSADR, 1997; CROSSA, 1990; YAN et al., 2000). Estes têm sido amplamente aplicados em programas de melhoramento de plantas, para a identificação de mega-ambientes, de genótipos ideais, para adaptação específica etc (não se uso vírgula antes de (etc.). Limitações inerentes a esses modelos (modelos de efeitos fixos), como a falta de flexibilidade para tratar dados desbalanceados e com heterogeneidade de variâncias, motivaram o desenvolvimento de métodos mais eficientes, como aqueles que utilizam modelos mistos para as análises. Piepho (1997, 1998) e Smith, Cullis e Thompson (2001) propuseram a análise MET sob a ótica de modelos mistos multivariados, utilizando estrutura de análise de fatores (AF) considerando ambientes/genótipos e interação como efeitos aleatórios.

Na literatura, esses modelos têm sido referidos frequentemente por modelos fator analíticos (FA) (BURGUEÑO et al., 2008; CROSSA et al., 2006; KELLY et al., 2007; SMITH; CULLIS; THOMPSON, 2001; SMITH et al., 2002) e têm mostrado grande versatilidade para a seleção de genótipos, uma vez que combinam o estudo da estabilidade e adaptabilidade em uma única abordagem. A vantagem dessa abordagem relaciona-se à capacidade do modelo em lidar com dados altamente desbalanceados, heterogeneidade de variâncias e covariâncias residuais e genotípicas. Esses modelos se destacam também por permitirem a inclusão de efeito dos resíduos heteroscedásticos de valores genéticos os quais constituem em um aspecto importante a ser considerado nas análises (HILL, 1984; RÖNNEGÅRD et al., 2010). Sabe-se

que a heterogeneidade de variâncias entre os genótipos é afetada pela heterogeneidade de variâncias entre os ambientes e vice-versa (EDWARD; JANNINK, 2006). Além do mais, esse modelo tem se mostrado útil para resumir o padrão de covariância em dados multivariados (HOGAN; TCHERNIS, 2004).

Não obstante, apesar das reconhecidas vantagens oferecidas pelos modelos FA, o método também possui limitações, como a necessidade da imposição de restrições e a não identificabilidade na estimação de parâmetros. Foi possível destacar ainda a dificuldade para construir intervalos de confiança para os componentes de variância, já que são aproximados e exigem suposições de normalidade assimptótica, além de uma grande demanda por recursos computacionais, bem como demanda por algoritmos mais eficientes que evitem, entre outros aspectos, a ocorrência em que a solução se encontra fora do espaço paramétrico-casos Heywood (MARDIA; KENT; BIBBY, 1979; MEYER, 2009).

Uma alternativa interessante, ao método frequentista é a utilização da inferência bayesiana. A análise bayesiana de modelos lineares mistos possibilita maior flexibilidade para a construção de intervalos de confiança para as estimativas das variáveis aleatórias, componentes de variância e efeitos fixos, já que todo processo de inferência é baseado na distribuição a posteriori. Essa flexibilidade do método bayesiano foi em parte ilustrada por Crossa et al. (2011) e Oliveira et al. (2015), que incorporaram regiões de credibilidade para os parâmetros bilineares que descrevem a interação no biplot do modelo AMMI. Sabe-se que a incorporação de inferência para escores genotípicos e ambientais, em modelos lineares-bilineares, apresenta grandes dificuldades aos métodos frequentistas paramétricos e não paramétricos (CROSSA et al., 2011; OLIVEIRA et al., 2015; YAN, 2010; YANG et al., 2009).

Segundo Sun et al. (1996), uma análise bayesiana pode ser obtida para os modelos de componentes de variância baseados na teoria da normalidade, permitindo, para qualquer parâmetro de interesse, uma detalhada inferência para amostras de tamanho finito. Além disso, informações adicionais podem ser incorporadas ao modelo sem maiores dificuldades, como é o caso de matrizes de correlações genéticas, o que torna os resultados mais precisos. Perez-Elizald, Jarquin e Crossa (2011), por exemplo, mostraram como informações históricas podem facilmente ser incorporadas na análise AMMI usando abordagem bayesiana. Recentemente, Jarquín et al. (2016) mostraram como essas informações poderiam ser incluídas no modelo SREG (*sites regression model*) usando abordagem bayesiana multinível (hierárquica).

Inovações metodológicas e aplicações da análise fatorial têm se desenvolvido rapidamente nos últimos anos, em parte devido ao maior acesso a ferramentas computacionais apropriadas. Em particular, podem-se destacar métodos iterativos de simulação, como o

Markov Chain Monte Carlo (MCMC), que abriram o acesso a tratamentos completamente Bayesianos dos modelos de análise de fatores (GEWEKE; ZHOU, 1996).

Uma proposta de análise bayesiana para o modelo FA foi apresentada por Campos e Gianola (2007). Nessa abordagem foram assumidas prioris como pressupostos básicos do modelo de análise de fatores padrão (clássica), o que evita a necessidade de imposição de restrições e reduz a demanda computacional que tem tornado o uso desses modelos pouco comum. Apesar da utilidade dessa proposta, ela apresenta algumas deficiências por não ser baseada no ajuste do modelo FA de Smith, Cullis e Thompson (2001), já que a proposta de Campos e Gianola (2007) só levou em consideração a modelagem da variância genética, ignorando a ambiental, além de amostrar primeiro os parâmetros do modelo misto e em seguida decompor os efeitos genéticos usando a análise bayesiana de fatores. Além disso, não foram consideradas situações em que se tenham dados perdidos ou em falta (nem todos os genótipos testados em todos ambientes), o que certamente constitui uma grande virtude dos modelos FA para análise de dados MET. As vantagens dos modelos FA para sintetizar os dados provenientes de programas de melhoramento, seja vegetal ou animal, podem ser encontradas em Burgueño et al. (2007, 2008), Kelly et al. (2007), Meyer (2009), Smith, Cullis e Thompson (2001, 2005), Smith et al. (2015), Stefanova e Buirchell (2010) e Tyrisevä et al. (2011).

Com este trabalho propõe-se um procedimento de inferência bayesiana para o modelo FA, com um processo de amostragem direta para os escores e cargas fatoriais, sem necessidade de rotacionalidade, visando a obtenção de uma estrutura simples, única e sem imposição de restrições dentro da estrutura do modelo misto, evitando a ocorrência de casos Heywood (variâncias negativas), além de o modelo ser identificável, ou seja, com todos parâmetros estimáveis.

1.1 Objetivos

i. Geral

Descrever formas gerais de como a heterogeneidade das covariâncias genéticas e residuais podem ser modeladas na perspectiva da análise fatorial em modelos mistos, utilizando-se decomposição espectral dos efeitos genéticos dentro da ótica Bayesiana.

ii. Específicos

a) Propor o uso da abordagem bayesiana para os modelos FA de Smith, Cullis e Thompson (2001) na análise de dados MET balanceados e desbalanceados;

b) Desenvolver um método de amostragem de cargas fatoriais que garanta a indentificabilidade e rotacionalidade do modelo sem necessidade de imposição de restrições;
c) Testar a robustez do método proposto na análise de dados desbalanceados usando dados simulados;
d) Selecionar o modelo com alta capacidade preditiva em dados MET faltantes.

2. REVISÃO DE LITERATURA

Nesta seção, são apresentados os principais aspectos teóricos relacionados à utilização do modelo FA (Fator Analítico). Na subseção 2.1, são apresentados os principais métodos para o estudo da interação entre genótipos e ambientes GEI. Na 2.2 é apresentado o procedimento de análise de fatores clássica. Na subseção 2.3, é apresentado o modelo fator analítico. Na subseção 2.4, apresenta-se a relação entre os modelos FA e SREG (*Site Regression*). Na subseção 2.5, são apresentados a relação entre os métodos de análise de componentes principais (PCA) e o modelo fator analítico (FA). Na subseção 2.6 apresentam-se os trabalhos realizados com análise FA usando modelos mistos, sob o enfoque frequentista, bem como os principais aspectos relacionados a esta abordagem. Na subseção 2.7 apresentam-se as principais estruturas de covariâncias usadas no estudo da interação GEI. Na subseção 2.8 são apresentados os principais conceitos relacionados à inferência bayesiana e na subseção 2.9 são apresentados os principais métodos usados para conduzir a análise de fatores bayesiana. Na última subseção são apresentados os principais estudos da interação GEI usando abordagem bayesiana.

2.1. Métodos para o estudo da interação genótipo x ambiente (GEI)

Vários modelos são comumente utilizados para descrever e interpretar a GEI em experimentos agrícolas, destacando-se entre os modelos propostos os lineares, os bilineares e os linear-bilineares. Deve se destacar que esses modelos envolvem métodos univariados e multivariados.

Dentre as propostas univariadas podemos citar a análise de variância conjunta de experimentos e os métodos de regressão linear (CROSSA, 1990; EBERHART; RUSSELL, 1966; FINLAY; WILKINSON, 1963). Contudo, os métodos multivariados têm se mostrado mais eficientes para explorar as informações contidas em conjuntos de dados MET. Dentre esses métodos podemos destacar a utilização da análise de componentes principais (PCA), a análise de agrupamento, a análise discriminante e a utilização de modelos lineares-bilineares como o modelo AMMI e o GGEbiplot (CROSSA, 1990; GAUCH, 2006; YAN et al., 2007). Esses últimos têm sido amplamente utilizados e se destacado na literatura por integrarem a análise univariada com o PCA, permitindo analisar a estabilidade e a adaptabilidade em uma única abordagem. Testes estatísticos e procedimentos não paramétricos podem ser utilizados para a seleção de componentes bilineares simplificando e melhorando a capacidade de predição dos modelos. Além disso, os padrões de respostas de genótipos e ambientes podem ser

visualizados graficamente usando biplots (GABRIEL, 1971), permitindo aos melhoristas a observação de genótipos de alto desempenho em determinadas regiões ou sub-regiões.

Crossa (2012) e Crossa e Cornelius (2002) destacam dois modelos, denominados respectivamente, de modelo de regressão fatorial (FR), dentro dos modelos lineares, e de regressão quadrados mínimos parciais (PLS) dentro da classe dos modelos bilineares. Esses modelos permitem a incorporação de covariáveis ambientais e genotípicas externas diretamente no modelo para mensurar as causas climáticas da GEI ou os fatores genéticos (marcadores moleculares) que influenciam a GEI. Tradicionalmente esses modelos foram propostos como de efeitos fixos e suas extensões como modelos para efeitos aleatórios ou mistos são mais recentes. A desvantagem da abordagem convencional é que o modelo tem pouca flexibilidade para tratar dados desbalanceados, além de o efeito genético ser assumido como independente do efeito ambiental e as variâncias residuais serem assumidas homogêneas, o que é pouco realista (CROSSA et al., 2006; EDWARDS; JANNINK, 2006; SMITH; CULLIS; THOMPSON, 2001). Outra crítica diz respeito ao método de visualização gráfica, que na maioria das abordagens não comporta qualquer incerteza sobre os escores plotados e que tem sido utilizado como critério para tomada de decisões (CROSSA et al., 2011; OLIVEIRA et al., 2015; YANG et al., 2009).

Essas deficiências têm motivado a busca por métodos mais flexíveis e eficientes. Entre os modelos testados no contexto frequentista, merecem especial destaque, os modelos mistos que fornecem uma maneira de lidar com as deficiências dos modelos AMMI, permitem modelar covariâncias genéticas heterogêneas, tendo em conta possíveis correlações entre os efeitos presentes no modelo (PIEPHO, 1997; SMITH; CULLIS; THOMPSON, 2001). Não obstante as vantagens citadas acima, com relação ao uso de modelos mistos, tem se verificado a tendência em assumir a mesma variância residual para todas as observações. A questão é que a variância residual entre ambientes existe e é importante que seja incluída em modelos tradicionais de avaliação de valores genéticos e considerada como heteroscedástica (HILL, 1984; RÖNNEGÅRD et al., 2010). Tais modelos, tendo variáveis explicativas representando resíduos heterocedásticos, são rotineiramente utilizados pelos melhoristas, mas as variáveis explanatórias são tipicamente não genéticas e a heterogeneidade genética pode estar presente e é incluída como efeito aleatório na variância residual do modelo.

Este fato levou Piepho (1997) a propor o uso de análise de fatores dentro de modelos mistos como forma de modelar a covariância genética. Tais modelos foram amplamente estudados por Smith, Cullis e Thompson (2001, 2005) e são denominados como fator analítico (FA). Essa abordagem combina, em um único modelo, metodologias de estudo de

adaptabilidade e estabilidade, além de permitir o ajuste de dados desbalanceados, modelagem da variância do erro e correlação espacial dentro do ambiente, além de lidar com a estrutura fator analítico (FA) a covariância genética para os ambientes que é a mais parcimoniosa e flexível que outras estruturas de covariância. Essas propriedades levaram a ampla aceitabilidade do modelo FA, bem como sua utilização para a análise MET em melhoramento de plantas (CROSSA et al., 2006; PIEPHO, 1997, 1998; SMITH; CULLIS; THOMPSON, 2001; YANG et al., 2009). Uma vez que os modelos mistos permitem que a informação de pedigree possa ser incorporada ao modelo FA para a GEI, obtêm-se um estimador mais acurado para o rendimento de genótipos (CROSSA, 2012; KELLY et al., 2009; OAKEY et al., 2007; SMITH; CULLIS; THOMPSON, 2005; SMITH et al., 2002). Uma vez que esses modelos incorporam o pedigree e consideram a heterogeneidade de covariância genética e residual, estes são considerados mais realistas na análise de dados MET para a modelagem da GEI. Contudo, a busca por modelos alternativos para estimar as variâncias genéticas e residuais heterogêneas tem despertado o interesse de vários pesquisadores no uso de ...nova abordagem, ou novas abordagens... como a bayesiana (EDWARDS; JANNINK, 2006) e *double hierarchical generalized linear models* (DHGLM) (RÖNNEGÅRD et al., 2010).

A crescente evolução tecnológica e computacional abriu caminho para o surgimento e disseminação de propostas baseadas na inferência bayesiana em estudos de GEI (BAUER et al., 2009; COTES et al., 2006; EDWARDS; JANNINK, 2006; THEOBALD; TALBOT; NABUGOOMU, 2002). Viele e Srinivasan (2000) mostraram como conduzir uma análise bayesiana para o modelo AMMI e, especialmente, como amostrar os vetores singulares genotípicos e ambientais, cujo suporte para a distribuição conjunta a posteriori não é trivial. Essa abordagem abriu caminho para outras contribuições (CROSSA et al., 2011; JARQUÍN et al., 2016; JOSSE et al., 2014; PEREZ-ELIZALDE; JARQUIN; CROSSA, 2011; OLIVEIRA et al., 2015; ORELLANA, 2012; SILVA et al., 2015). As vantagens da abordagem bayesiana são muito sugestivas e permitem superar limitações dos modelos tradicionais. Além disso, com o uso de novos métodos de amostragem e aumento da velocidade computacional, uma ampla gama de ferramentas Bayesianas está à disposição dos pesquisadores.

Como já especificado, este trabalho pretende explorar as propriedades dos modelos FA sob a ótica bayesiana, que permite entre outras coisas maior flexibilidade para adicionar informações ao modelo, obtendo-se uma estimação mais acurada de parâmetros. Sendo esses modelos (FA-AMMI e FA-SREG) os mais usados atualmente na análise de dados MET e a metodologia bayesiana flexível e com inúmeras vantagens, a extensão dessa classe de modelos

via abordagem bayesiana torna-se uma prática imprescindível. A seguir apresentaremos os principais métodos que veêm norteando o trabalho dos melhoristas.

2.1.1 O modelo básico de análise da GEI

O modelo linear de base bidirecional de efeitos fixos para análises GEI considera que a resposta média empírica, $\bar{y}_{ij}$, do genótipo i (i = 1,2, ..., g) no ambiente j (j = 1,2, ... , s) com r repetições em cada uma das células da matriz g × s é expressa como:

$$\bar{y}_{ij} = \mu + \tau_i + \delta_j + (\tau\delta)_{ij} + \bar{\varepsilon}_{ij} \qquad (1)$$

em que μ é a média geral sobre todos os genótipos e ambientes, τ_i é o efeito principal do genótipo, δ_j é o efeito principal do j-ésimo ambiente, $(\tau\delta)_{ij}$ é o efeito da interação (GE) do genótipo i, no ambiente de ordem j, e $\bar{\varepsilon}_{ij}$ é o erro médio, assumido ser Normal Identicamente Distribuído (NID) $(0, \sigma_\varepsilon^2 / r)$ sendo σ_ε^2 a variância de erro dentro do ambiente, e o símbolo r representa o número de observações por célula. Para um modelo aleatório completo, assume-se que τ_i, δ_j, e $(\tau\delta)_{ij}$ são normalmente distribuídos e independentemente, com variâncias σ_τ^2, $\sigma_\delta^2, \sigma_{\tau\delta}^2$, respectivamente. Adicionar os efeitos ao modelo (1) como blocos ao acaso, ou qualquer tipo de delineamento de blocos incompletos, não representa um grande problema. Além disso, a modelagem dos resíduos por meio de análises espaciais não apresenta mais dificuldade e é uma prática que deve ser utilizada rotineiramente em qualquer experimento de campo (CROSSA, 2012).

Yates e Cochran (1938) introduziram um modelo em que o termo GEI está linearmente relacionado com o efeito principal ambiental $(\tau\delta)_{ij} = \partial_i \bar{y}_{.j}$, de tal forma que o parâmetro de estabilidade ∂_i é a regressão do desempenho do genótipo sobre o ambiente de média $\bar{y}_{.j}$. Isto foi, formalmente, apresentado, mais tarde por Finlay e Wilkinson (1963) e estendido por Eberhart e Russell (1969) para incluir o desvio de regressão como outro parâmetro de estabilidade (embora isto seja de fato o ajuste do modelo) (CROSSA, 2012).

2.1.2 Modelos lineares-bilineares-MLB de efeitos fixos

Os modelos linear-lilineares (MLB) de efeitos fixos são descritos por Williams (1952) como:

$$\bar{y}_{ij} = \mu + \tau_i + \lambda\alpha_i\gamma_j + \bar{\varepsilon}_{ij}, \quad (2)$$

em que λ é o maior valor singular de $\mathbf{ZZ}^{\mathrm{T}}$ ou $\mathbf{Z}^{\mathrm{T}}\mathbf{Z}$ (para $\mathbf{Z} = \bar{y}_{ij} - \bar{y}_{i.}$) e α_i e γ_j são os autovetores correspondentes. Gollob (1968) e Mandel (1969) estenderam o trabalho de Williams (1952), considerando o termo bilinear GEI como $(\tau\delta)_{ij} = \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk}$ Assim, a formulação geral do modelo linear-bilinear passou a ser:

$$\bar{y}_{ij} = \mu + \tau_i + \delta_j + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \bar{\varepsilon}_{ij}, \quad (3)$$

em que o λ_k é o valor singular do k-ésimo componente bilinear (multiplicativo) seguindo a ordem $\lambda_1 \geq \lambda_2 ... \geq \lambda_t$; α_{ik} são elementos dos k-ésimos vetores singulares esquerdos da verdadeira interação e representa a sensibilidade genotípica de fatores ambientais hipotéticos representados pelo k-ésimo vetor singular direito com elementos γ_{jk}. Os α_{ik} e γ_{jk} satisfazem as restrições $\sum_{k=1}^{g} \alpha_{ik}\alpha_{ik'} = \sum_{k=1}^{s} \gamma_{jk}\gamma_{jk'} = 0$ para $k \neq k'$ e $\sum_i \alpha_{ik}^2 = \sum_j \gamma_{jk}^2 = 1$. Gabriel (1978) descreveu o ajuste por mínimos quadrados da Eq. 3 e explicou como o termo GE, $\mathbf{Z} = \bar{y}_{ij} - \bar{y}_{i.} - \bar{y}_{.j} + \bar{y}_{..}$, é submetido à decomposição em valores singulares (SVD) após o ajuste para os termos aditivos (lineares). Gauch (1988) chamou Eq. 3 de modelo de efeitos principais aditivos e interação multiplicativa (AMMI).

Outras classes de modelos linear-bilineares descritos por Cornelius, Crossa e Seyedsadr (1996) são: Modelo de regressão de genótipos (**GREG**)

$$\bar{y}_{ij} = \mu_i + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \bar{\varepsilon}_{ij}; \quad (4)$$

em que: μ_i é a média do i-ésimo genótipo, e as definições dos parâmetros λ_k, α_{ik}, e γ_{jk} são semelhantes às suas no modelo AMMI. No entanto, no modelo SREG os principais efeitos do ambiente são absorvidos em termos bilineares.

Modelo de Regressão por locais (ambientes) (**SREG**)

$$\bar{y}_{ij} = \mu_j + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \bar{\varepsilon}_{ij}; \quad (5)$$

em que: μ_j é a média do j-ésimo ambiente e as definições dos parâmetros λ_k, α_{ik}, e γ_{jk} são semelhantes às suas no modelo AMMI. No entanto, no modelo SREG os principais efeitos do genótipo são absorvidos em termos bilineares. O modelo SREG tem sido usado para o agrupamento de ambientes sem mudança estatisticamente significativa genotípica classificação (CROSSA; CORNELIUS, 1997; CROSSA; YANG; CORNELIUS, 2004).

Modelo Completamente Multiplicativo (**COMM**), obtido pela omissão de todos os termos aditivos

$$\bar{y}_{ij} = \mu + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \bar{\varepsilon}_{ij}; \quad (6)$$

Modelo multiplicativo Shifted (**SHMM**)

$$\bar{y}_{ij} = \beta + \sum_{k=1}^{t} \lambda_k \alpha_{ik} \gamma_{jk} + \bar{\varepsilon}_{ij}; \quad (7)$$

em que: β é o parâmetro Shifted, e as definições dos parâmetros λ_k, α_{ik}, e γ_{jk} são semelhantes às suas no modelo AMMI.

As estimativas de mínimos quadrados dos efeitos aditivos destes modelos, bem como os elementos residuas da matriz **Z** são dados como:

AMMI $\hat{\mu} = \bar{y}_{..}$, $\hat{\tau} = \bar{y}_{\text{i.}} - \bar{y}_{..}$, $\hat{\delta} = \bar{y}_{\text{.j}} - \bar{y}_{..}$, $z_{ij} = \bar{y}_{ij} - \bar{y}_{\text{i.}} - \bar{y}_{\text{.j}} + \bar{y}_{..}$

GREG $\hat{\mu}_i = \bar{y}_{\text{i.}}$, $z_{ij} = \bar{y}_{ij} - \bar{y}_{\text{i.}}$

SREG $\hat{\mu}_j = \bar{y}_{\text{.j}}$, $z_{ij} = \bar{y}_{ij} - \bar{y}_{\text{.j}}$

COMM $z_{ij} = \bar{y}_{ij}$

SHMM $z_{ij} = \bar{y}_{ij} - \hat{\beta}, \hat{\beta} = \bar{y}_{..} - \sum_{k=1}^{m} \lambda_k \bar{\alpha}_k \bar{\hat{\gamma}}_k$

Assim, a adição de qualquer termo na regressão linear sobre covariáveis, são casos particulares do modelo linear-bilinear geral-GLBM.

Na notação matricial, esses modelos lineares-bilineares podem ser expressos como

$$\mathbf{Y} = \sum_{k=1}^{t} \beta_k \mathbf{X}_k + \mathbf{A}\mathbf{\Lambda}\mathbf{G}^{\mathrm{T}} + \mathbf{E} \quad (8)$$

em que $\mathbf{Y} = \left[\bar{y}_{ij} \right], \mathbf{X}_k = \left[x_{kij} \right], \mathbf{\Lambda} = diag\left(\lambda_k, k = 1, 2, ..t \right), \lambda_1 \geq \lambda_2 ... \geq \lambda_t, \mathbf{A} = \left(\alpha_1, ..., \alpha_t \right),$ $\mathbf{G} = \left(\gamma_1, ..., \gamma_2 \right)$ e $\mathbf{A}^{\mathrm{T}}\mathbf{A} = \mathbf{G}^{\mathrm{T}}\mathbf{G} = \mathbf{I}_t$. O x_{kij} são constantes conhecidas e $\beta_k, \lambda_k, \alpha_{ik}$, e γ_{ik} são parâmetros a serem estimados (CROSSA, 2012; CROSSA; CORNELIUS, 2002; CORNELIUS; SEYEDSADR, 1996). É possível estudar a GEI ajustando aproximações dimensionais baixas das medias por meio de decomposição em valores singulares (SVD) da estrutura presente na tabela de dupla entrada.

O modelo linear-bilinear GREG definido acima é uma reparametrização do modelo de análise de estabilidade de Finlay e Wilkinson (1963) e dos modelos Eberhart e Russell (1969) que executam as regressões lineares de genótipo sobre as médias dos ambientes. No modelo GREG, o primeiro termo multiplicativo, $\lambda_1 \alpha_{i1} \gamma_{j1}$ é percebido como as regressões dos genótipos, com coeficientes α_{i1} em índices ambientais (ou) índice ambiental e γ_{j1} (o parâmetro de escala λ_1 pode ser absorvido por α_{i1} ou γ_{j1}, ou parcialmente em cada um), e o desvio modelado como componentes multiplicativos, desde que $t > 1$ (CROSSA, 2012; CROSSA; CORNELIUS, 2002; CORNELIUS; SEYEDSADR, 1997).

Crossa (2012) destaca a existência de várias razões estatísticas, bem como biológicas para preferir SREG ao AMMI para avaliar a GEI, tais como: (1) para o mesmo número de termos bilineares, SREG é um modelo mais parcimonioso do que AMMI; (2) SREG incorpora o efeito principal de genótipos (G) diretamente para a análise estatística da GEI, isto é, ambos os efeitos G e GEI (GGE) são combinados e estimados em conjunto; isso é importante para atingir os objetivos de melhoristas, que requerem a inclusão do efeito principal dos genótipos no modelo; (3) O modelo SREG misto pode ser montado muito mais facilmente do que o modelo AMMI misto; e (4) o modelo SREG misto, tal como comprovado por Burgueño et al. (2008), é útil para delinear mega-ambientes usando uma abordagem estatística formal, com base no modelo fator analítico (FA). Uma variação do modelo SREG foi apresentada por Yan et al. (2000) e denominada por GGEbiplot que utiliza apenas dois termos bilineares sob a justificativa que os demais eixos não têm interpretação biológica ou agronômica. O GGEbiplot se tornou muito popular e com ampla aplicação no contexto do melhoramento de plantas por permitir entre outros aspectos a identificação do genótipo ideal e a específicação de mega ambientes. Méritos, limitações e comparações dos modelos AMMI e GGEbiplot podem ser encontrados na literatura, como por exemplo, em Crossa et al. (2006), Gauch, Piepho e Annicchiarico (2008), Yan et al. (2007) e Yang et al. (2009).

2.1.3 Modelos linear-bilineares de efeitos mistos (MLBM)

Piepho (1997) propôs um modelo misto considerando genótipos como efeitos fixos e ambientes como aleatórios o que leva à estimação de uma matriz de covariância dos efeitos genéticos em ambientes. Essa proposta não foi muito bem aceita e Smith, Cullis e Thompson (2001) propuseram que os efeitos de genótipos fossem aleatórios e dos ambientes fossem fixos, uma vez que assumir efeitos de genótipos como fixo implicaria em herdabilidade 1 e não se teria a predição dos seus efeitos, que é o objetivo dos programas de melhoramento.

O modelo linear misto básico utilizado para ajuste dos dados de g genótipos, s locais, e r repetições na busca de subconjuntos de ambientes ou genótipos é

$$\mathbf{Y} = \mathbf{Xb} + \mathbf{Z_r r} + \mathbf{Z_g g} + \mathbf{e} \qquad (9)$$

sendo $\mathbf{X}$ a matriz de incidência dos efeitos fixos de ambientes e $\mathbf{Z_r}$ e $\mathbf{Z_g}$ são as matrizes de incidência dos efeitos aleatórios de repetições dentro de ambientes e de genótipos, respectivamente. O efeito aleatório de genótipos em ambientes combina os efeitos principais de genótipos e GEI (GGE). O vetor $\mathbf{b}$ denota os efeitos fixos de ambientes; os vetores $\mathbf{r}$, $\mathbf{g}$, e $\mathbf{e}$ contêm efeitos aleatórios de repetições dentro de ambientes, genótipos dentro de ambientes e resíduos dentro de ambientes, respectivamente, e estão a ser assumidos como aleatórios e normalmente distribuídos com vetores de médias zero e matrizes de variâncias e covariâncias $\mathbf{E}$, $\mathbf{G}$ e $\mathbf{R}$, respectivamente. As matrizes de variâncias e covariâncias $\mathbf{R}$ e $\mathbf{E}$ são assumidos ter a estrutura simples de componente de variância, ou seja, $\mathbf{E} = \left(\sigma^2_{r_j}, j = 1..., \mathrm{p}\right) \otimes \mathbf{I}_r$ e $\mathbf{R} = \sigma^2_e \otimes \mathbf{I}_{rgp}$, em que as matrizes de identidade $\mathbf{I}_r$ e $\mathbf{I}_{rgs}$ são de ordem r e $r \times g \times s$, respectivamente, $\sigma^2_{r_j}$, σ^2_e e são as repetições dentro do ambiente de ordem j e variâncias residuais, respectivamente, e $\otimes$ é o de produto de Kronecker (ou direto) das duas matrizes. A estrutura de $\mathbf{R}$ assume que os resíduos das parcelas de campo em cada ambiente (ou seja, os elementos do vetor $\mathbf{e}$) não são espacialmente correlacionados. No entanto, quando a informação de campo está disponível, a abordagem de modelo espacial usando tais modelos como o processo autorregressivo bidimensional, na direção de linhas e colunas no campo pode ser incorporada nas análises. A solução ($\hat{\mathbf{b}}$) para o vetor de médias de efeitos fixos de ambiente e os vetores de efeitos aleatórios ($\hat{\mathbf{r}}$ e $\hat{\mathbf{g}}$) é obtida a partir das equações do modelo misto (BURGUEÑO et al., 2008; CROSSA, 2012; KELLY et al., 2009; OAKEY et al., 2007).

A matriz de variância e covariância $\mathbf{G}$ é indexada por dois fatores, ambientes e genótipos, e, portanto, pode ser escrita como o produto de Kronecker de duas matrizes que indexam esses fatores, $\mathbf{G} = \Sigma_g \otimes \mathbf{I}_g$, onde o j-ésimo elemento da diagonal $s \times s$ da matriz Σ_g é o $\sigma^2_{g_j}$ da variância genética dentro do j-ésimo ambiente, e o jj'-ésimo elemento fora da diagonal é a $\rho_{jj}\sigma_{g_j}\sigma_{g_{j'}}$ covariância genética "entre os ambientes j e j'; assim $\rho_{jj'}$ é a correlação dos efeitos genéticos entre ambientes j e j'. Como para o fator genotípico, a matriz identidade de $\mathbf{I}_g$ (de ordem g) é utilizado quando se assume que os genótipos não estão relacionados, e o valor de genético de cada um dos genótipos só irá ser predito pelo valor das respostas empíricas do próprio genótipo. A componente ambiental da $\mathbf{G}$, Σ_g, pode ser modelado pela FA, enquanto a componente genotípica de $\mathbf{G}$ é modelado pela matriz identidade $\mathbf{I}_g$, que não assume qualquer relação entre os genótipos. No entanto, se existe uma matriz $\mathbf{A}$ de parentesco, utilizando o coeficiente de parentesco entre os genótipos, esta pode ser utilizada (CROSSA, 2012).

2.2. Análise fatorial (AF)

A análise fatorial (AF) é uma técnica de análise multivariada que descreve as relações de covariâncias entre muitas variáveis em termos de algumas variáveis não observáveis, ou seja, "fatores latentes" hipoteticamente existentes. Tais "fatores" não podem ser observados diretamente, em vez disso, eles são inferidos a partir de variáveis diretamente medidas (mensuradas), ...por meio ... da análise da estrutura de covariâncias das variáveis observadas.

A formulação comum da análise fatorial assume que as variáveis mensuráveis, como os escores, são manifestações de uma construção latente subjacente. A formulação de variável latente pode ser útil para a redução de dados, isto é, sintetizando observações multivariadas usando uma variável de dimensionalidade menor. Discussões e revisão completa foram apresentadas por Bartholomew e Knott (1999). Geweke e Zhou (1996), Press e Shigemasu (1997) e Rowe (2002, 2003) discutiram a necessidade de implantação bayesiana da análise de fatores.

O modelo k-fatorial pode ser escrito da seguinte maneira:

$$\mathbf{y} = \mathbf{\Gamma}_k \mathbf{f} + \boldsymbol{\mu} + \mathbf{e}, \tag{10}$$

em que $\mathbf{y}(p \times 1)$ é um vetor aleatório, $\mathbf{\Gamma}(p \times k)$ é a matriz de cargas com τ_{ij} carregamento da i-ésima variável no j-ésimo fator, $\mathbf{f}(\mathrm{k} \times 1)$ é o vetor aleatório dos fatores latentes (comuns),

$\boldsymbol{\mu}(p\times 1)$ é o vetor de médias de $\mathbf{y}$, e $\mathbf{e}(p\times 1)$ é o termo erro, também conhecido como fator único. A análise de fatores assume que: $E(\mathbf{f})=0$, $Var(\mathbf{f})=\mathbf{I}$, $E(\mathbf{e})=0$, $Cov(\mathbf{e}_k,\mathbf{e}_l)=0$ (com $k\neq l$), $Var(\mathbf{e})=\boldsymbol{\Psi}=diag(\psi_1,\psi_2,...,\psi_p)$ e $Cov(\mathbf{e},\mathbf{f})=0$.

No modelo AF, $\mathbf{y}$ é mensurável e, portanto, observável. Se a maior parte da variação de qualquer subgrupo de $\mathbf{y}$ for causada por um fator comum, então este subgrupo de variáveis manifesta-se como uma variável e pode ser reduzido essencialmente a uma dimensão.

O modelo FA é análogo a um modelo de regressão linear (M.R.L.), exceto que $\mathbf{f}$ são variáveis independentes não observáveis. O "erro" na regressão linear é muitas vezes referido como flutuações devido à medição em diferentes indivíduos e supõe aleatoriedade. As cargas fatoriais atuam como coeficientes de regressão em M.R.L. Portanto τ_{ij} representam como i-ésima variável se relaciona para o j-ésimo fator.

A porção da variância da j-ésima variável de interesse devido aos fatores comuns é chamada de i-ésima comunalidades e é dado por:

$$h_i^2=\sum_{j=1}^{q}\tau_{ij}^2$$

sendo que $\sigma_{ii}=\text{var}(y_i)=h_i^2+\psi_i$. A parte restante da variância da i-ésima variável é chamada de variância única ou específica, simplesmente porque, essa parte da variância não é uma contribuição de todos os fatores comuns. A variância específica é igual a i-ésima variância residual, nomeadamente:

$$\text{var}\left(y_i-\sum_{j=1}^{q}\tau_{ij}f_i\right)=\text{var}\left(e_i\right)=\psi_i=\sigma_{ii}-h_i^2.$$

O modelo (10) implica que a matriz de variâncias e covariâncias (doravante será tratada como matriz de covariâncias (MC)) de $\mathbf{y}$ pode ser decomposta como:

$$\text{var}\left(\mathbf{y}\right)=\boldsymbol{\Sigma}=\boldsymbol{\Gamma}_k\boldsymbol{\Gamma}_k^{\mathrm{T}}+\boldsymbol{\Psi} \qquad (11)$$

e $\text{cov}\left(y_i,y_j\right)=\sum_{k=1}^{q}\tau_{ik}\tau_{jk}, i\neq j$. Nota-se que as covariâncias não dependem de $\boldsymbol{\Psi}$. Isto faz sentido para explicar como o modelo FA é usado para explorar o padrão fatorial via aprendizagem sobre interdependência entre as varáveis de interesse. Note que estas covariâncias não são independentes da escala em que as variáveis são medidas, e por isso é recomendado padronizar as variáveis ou trabalhar com as correlações. Com o modelo (10), no

entanto as correlações dependem tanto de $\mathbf{\Gamma}$ e $\mathbf{\Psi}$. Assim, podemos trabalhar com uma parametrização diferente do modelo AF dado a seguir:

$$D^{-\frac{1}{2}} = (\mathbf{y} - \boldsymbol{\mu}) = \mathbf{\Gamma}_k \mathbf{f} + \mathbf{e}, \tag{12}$$

Em que $\mathbf{D} = diag(\mathbf{\Sigma})$, $\boldsymbol{\mu}$, $\mathbf{f}$ e $\mathbf{e}$ foram definidos anteriormente. Temos então $\mathbf{R} = corr(y) = \mathbf{D}^{-\frac{1}{2}} \mathbf{\Sigma} \mathbf{D}^{-\frac{1}{2}} = \mathbf{\Gamma}_k \mathbf{\Gamma}_k^{\mathrm{T}} + \mathbf{\Psi}$ que implica imediatamente em τ_{ij} satisfazendo $h_i^2 = \sum_{j=1}^{q} \tau_{ij}^2 \leq 1$ e $\psi_i = 1 - h_i^2$. Note que h_i^2 é agora a proporção da variância da i-ésima variável padronizada explicada pelos fatores comuns e ψ_i é a i-ésima proporção residual não explicada da variância. Assim:

$$\mathbf{R} = corr(\mathbf{y}) = \mathbf{D}^{-\frac{1}{2}} \mathbf{\Sigma} \mathbf{D}^{-\frac{1}{2}} = \mathbf{\Gamma}_k \mathbf{\Gamma}_k^{\mathrm{T}} + \mathbf{\Psi}(\Gamma_k) \tag{13}$$

em que: $\mathbf{\Psi} = diag(1 - h_1^2, ..., 1 - h_p^2)$. Portanto, $\rho_{ij} = corr(y_1, y_j) = \sum_{j=1}^{q} \frac{\tau_{il}}{\sqrt{1 - h_i^2}} \frac{\tau_{jl}}{\sqrt{1 - h_j^2}}$ e as correlações dependem somente de $\mathbf{\Gamma}$.

É mais sensato parametrizar as matrizes de correlação, em vez de matrizes de covariância, porque o padrão dos fatores não depende das escalas das variáveis de interesse. Em uma AF, há vários fatores a serem considerados, como discutido em Anderson e Rubin (1955), Cao (2010) e Rowe (1998, 2003), a saber: i) existência do modelo, ii) identificabilidade, iii) determinação da estrutura, iv) estimativa dos parâmetros, v) teste de hipóteses, vi) determinação do número de fatores e vii) estimativa dos escores fatoriais.

2.2.1 Relevância do modelo

O problema fundamental na AF é identificar se é possível explicar relações entre variáveis observadas por menos fatores latentes.

Na AF, sabemos que a escolha de um modelo específico depende da escolha que se faz para q. Suponhamos por exemplo que se tenha $k = p$ (número de observações p igual ao número de cargas q), temos $\mathbf{\Gamma}_p = \mathbf{\Sigma}^{\frac{1}{2}}$ e $\mathbf{\Psi} = 0$ e (11) sustenta isto. O que nos resta a fazer é verificar todos os pressupostos da distribuição que podemos fazer sobre, $\mathbf{f}$ e $\mathbf{e}$. Da mesma forma

em (12) podemos ter $\mathbf{\Gamma}_p = R^{\frac{1}{2}}$ para manter (13). Ignorando a distribuição podemos afirmar que o problema básico da AF é determinar o número mínimo de elementos k para que as equações (11) e (13) tenham, solução. Para $k < p$ e $\mathbf{\Sigma}$, dado que deve se determinar se existe uma solução $(\mathbf{\Gamma}_k, \mathbf{\Psi})$ para $\sigma_{ii} = \psi_{ii} + \sum_{k=1}^{q} \tau_{ik}^2$ e $\sigma_{ij} = \sum_{k=1}^{q} \tau_{ik}\tau_{jk}, i \neq j$. Em (11) com $\sum_{k=1}^{q} \tau_{ij}^2 \leq 1$ para (10).

Vamos considerar que o modelo (11) tem $p(p+1)/2$ elementos de $\mathbf{\Sigma}$, p elementos da diagonal de $\mathbf{\Psi}$ e a matriz de cargas $\mathbf{\Gamma}_q$ tem pk elementos. Além disso, qualquer $\mathbf{\Gamma}_k$ pode ser substituída por $\mathbf{\Gamma}_k\mathbf{Q}$, em que $\mathbf{Q}$ é a matriz $k \times k$ ortogonal com $\mathrm{k}(\mathrm{k}-1)/2$ elementos independentes, desde que $(\mathbf{\Gamma}_k, \mathbf{\Psi})$ seja a solução, assim $(\mathbf{\Gamma}_k\mathbf{Q}, \mathbf{\Psi})$. Para ver esta falta de indentificabilidade, seja $\mathbf{Q}$ uma matriz $k \times k$ ortogonal , $\mathbf{\Gamma}_k^* = \mathbf{\Gamma}_k\mathbf{Q}$ e $\mathbf{f}^* = \mathbf{Q}^{\mathrm{T}}\mathbf{f}$ então (10) torna-se $\mathbf{y} = \mathbf{\Gamma}_k^*\mathbf{f}^* + \boldsymbol{\mu} + \mathbf{e} = \mathbf{\Gamma}_k\mathbf{Q}\mathbf{Q}^{\mathrm{T}}\mathbf{f} + \boldsymbol{\mu} + \mathbf{e} = \mathbf{\Gamma}_k + \boldsymbol{\mu} + \mathbf{e}$ sendo $\mathbf{Q}\mathbf{Q}^{\mathrm{T}} = \mathbf{I}$ e $\mathbf{\Gamma}_k^*\left(\mathbf{\Gamma}_k^*\right)^{\mathrm{T}} + \mathbf{\Psi} = \mathbf{\Gamma}_k\left(\mathbf{Q}\mathbf{Q}^{\mathrm{T}}\right)\mathbf{\Gamma}_k^{\mathrm{T}} + \mathbf{\Psi} = \mathbf{\Gamma}_k\mathbf{\Gamma}_k^{\mathrm{T}} + \mathbf{\Psi} = \mathbf{\Sigma}$. Portanto, teremos somente $pk - k(\mathrm{k}-1)/2$ graus de liberdade (g.l.) em $\mathbf{\Gamma}_k$, assim os graus de liberdade à direita de (11) serão $pk - k(\mathrm{k}-1)/2 + p$. Assim, considerando a diferença entre os graus de liberdade entre aqueles $\mathbf{\Sigma}$ $(p(p+1)/2)$ e os da direita de (11) temos:

$$\frac{1}{2}p(p+1) - \frac{1}{2}\mathrm{k}(\mathrm{k}+1) - p - pk = \frac{1}{2}\left[(p-k)^2 - (p+k)\right] \qquad (14)$$

Se (14) for menor ou igual ao zero, espera-se que exista pelo menos uma solução para $\mathbf{\Sigma}$ arbitrário. Caso (14) seja maior que zero, esperamos somente uma única solução para algumas matrizes $\mathbf{\Sigma}$, do modelo fatorial oferecendo uma estrutura mais simples que a matriz de covariâncias completas. Em (13) a situação é similar para $\mathbf{R}$ com $p(p-1)/2$ g.l. $\mathbf{\Gamma}_q$ tem $pk - k(\mathrm{k}-1)/2$ g.l. e teremos:

$$\frac{1}{2}p(p-1) - \frac{1}{2}q(q+1) - pq = \frac{1}{2}\left[(p-q)^2 - (p+q)\right]\text{g.l.}$$

como dados em (14) para (13).

Um exemplo ilustrativo

Em análise de dados MET é comum retermos apenas duas cargas fatoriais para explicar o padrão dos dados da interação genótipo x ambiente (GEI). Assim sendo, vamos supor que $p = 2$ e resolvendo (11) e (13) para $\mathbf{\Gamma}_q$ com $k = 0,1,2$ observa-se quando $k = 0$, têm-se $\mathbf{\Gamma}_0$ e somente temos uma única solução quando y_i são variáveis independentes.

Quando $k = 1$, nós precisamos resolver o sistema de equações Para $\tau_{11}, \tau_{21},$ e ψ_2:

$$\begin{cases} \sigma_{11} = \psi_1 + \tau_{11}^2 \\ \sigma_{22} = \psi_2 + \tau_{21}^2 \\ \sigma_{21} = \tau_{11}\tau_{21} \end{cases}$$

Note que (14) é menor que zero neste caso. Portanto, dado um $\mathbf{\Sigma}$ arbitrário, qualquer Γ_1 satisfaz $\tau_{11}\tau_{21} = \sigma_{21}$ produzindo a solução para Ψ por $\psi_1 = \sigma_{11} - \tau_{11}^2$ e $\psi_2 = \sigma_{22} - \tau_{21}^2$, dado $\psi_1, \psi_2 \geq 0$.

Seja $\tau_{11}^* = \tau_{11} / \sqrt{\sigma_{11}}$ e $\tau_{21}^* = \tau_{21} / \sqrt{\sigma_{22}}$ e uma equação equivalente em

$\tau_{11}\tau_{21} = \sigma_{21}$

Dada por

$\tau_{11}\tau_{21} = \sigma_{21}$,

isto é

$$\tau_{11}^*\tau_{21}^* = \rho_{21} \qquad (15)$$

E verifica-se que $\tau_{11}^* = \sqrt{|\rho_{12}|}$, $\tau_{22}^* = sign(\rho_{12})\sqrt{|\rho_{12}|}$ satisfazendo (15), e as soluções são dadas por:

$\tau_{11} = \sqrt{|\rho_{12}|}$,

$\tau_{22} = sign(\rho_{12})\sqrt{\sigma_{22}}\sqrt{|\rho_{12}|}$,de fato $\tau_{11}^2 = \sigma_{11}|\rho_{12}| \leq \sigma_{22}$, $\tau_{12}^2 = \sigma_{22}|\rho_{12}| \leq \sigma_{22}$ se $\rho_{12} \leq 1$e $\psi = \sigma_{11} - \tau_{11}^2 = \sigma_{11}\left(1 - |\rho_{12}|\right)$ e $\psi = \sigma_{22} - \tau_{21}^2 = \sigma_{22}\left(1 - |\rho_{12}|\right)$.

Além disso, nós temos que a família das soluções

$\tau_{11}^* = a\sqrt{|\rho_{12}|}$ e $\tau_{22}^* = a^{-1}sign(\rho_{12})\sqrt{\sigma_{22}}\sqrt{|\rho_{12}|}$,

assim $\tau_{11} = a\sqrt{|\rho_{12}|}$ e $\tau_{22} = a^{-1} sign(\rho_{12})\sqrt{\sigma_{22}}\sqrt{|\rho_{12}|}$

para todo **a** desde que

$a^2(\rho_{12}) \leq 1$ e $a^{-2}(\rho_{12}) \leq 1$

ou equivalente satisfazendo

$|\rho_{12}| \leq a^2 \leq |\rho_{12}|^{-1}$ com $\psi_1 = \sigma_{11}\left(1-a^2\right)|\rho_{12}|$ e $\psi_2 = \sigma_{22}\left(1-a^2\right)|\rho_{12}|$.

Quando $|\rho_{12}| = 0$, se não τ_{11} e τ_{22} ficam iguais a zero.

Como anteriormente mencionado para $k = 2$, existe sempre uma solução tomando $\mathbf{\Gamma}_2 = \mathbf{\Sigma}^{\frac{1}{2}}$ e $\mathbf{\Psi} = \mathbf{0}$. De fato teremos de novo infinitas soluções e neste caso dadas por $\mathbf{\Sigma}^{\frac{1}{2}}\mathbf{Q}$ para qualquer matriz ortogonal $\mathbf{Q} \in IR^{2x2}$.

Podemos determinar $\mathbf{\Gamma}_k$ para $k = 0,1,2$ do modelo (12) resolvendo (13). Quando $k = 0$ temos $\mathbf{\Gamma}_0$ é a solução única quando os y_i são variáveis aleatórias independentes. Quando $k = 1$, então a análise anterior aplica-se com $\sigma_{11}\sigma_{22} = 1$. Quando $k = 2$, existe solução dada por $\mathbf{\Gamma}_2 = R^{\frac{1}{2}}$ e $\mathbf{RQ}$ que é a solução para toda a matriz ortogonal $\mathbf{Q} \in IR^{2x2}$.

2.2.3 Princípio de análise fatorial (AF)

Conforme Johnson e Wichern (2007), o método de fator principal se comporta como uma modificação do método do componente principal.

A análise fatorial (AF) e a análise de componentes principais (PCA) são usadas para a mesma finalidade, ou seja, expressar a informação contida numa série de observações em um menor número de dimensões.

Por ambas as técnicas terem a mesma finalidade elas são comuns em alguns aspetos. Ambas têm como objetivo explicar os dados com dimensão reduzida. A PCA tenta explicar o total de variâncias dos dados possíveis e é composto por uma transformação linear das variáveis originais. Por outro lado, a AF tenta responder pelas covariâncias entre as variáveis originais e transformá-las em variáveis latentes.

O princípio da AF segue a lógica dos componentes principais e muitas vezes servem como um procedimento *adhoc* para determinar o número de fatores.

A AF via PCA é usado quando os modelos (10) e (12) estão bem definidos ,ou seja, quando (14) for maior que zero. É comum se trabalhar com variáveis padronizadas para evitar problemas de escala. Primeiro começa-se com algumas estimativas preliminares das comunalidades h_i^2, assim os elementos da diagonal de $\mathbf{R}-\mathbf{\Psi}$ são estimados por $\hat{h}_i^2$. Então $\mathbf{R}-\hat{\mathbf{\Psi}}$ pode ser reescrito a partir do teorema de decomposição espectral como:

$$\mathbf{R}-\hat{\mathbf{\Psi}}=\sum_{i=1}^{p}\eta_i\boldsymbol{\alpha}_i\boldsymbol{\alpha}_i^t \tag{16}$$

em que $\eta_i^{'}s$ são autovalores com $\eta_1\geq\eta_2...\geq\eta_p$ de $\mathbf{R}-\hat{\mathbf{\Psi}}$ e $\boldsymbol{\alpha}_i^{'}s$ são autovetores correspondentes. Uma escolha, $k(<p)$ de autovalores não-negativos e estimativas de $\mathbf{\Gamma}_k$ dados por:

$$\hat{\mathbf{\Gamma}}_k=V\Lambda^{\frac{1}{2}}$$

em que: $\mathbf{\Lambda}=diag(\eta_1,...\eta_k)$ e $\mathbf{V}=(\boldsymbol{\alpha}_1,...\boldsymbol{\alpha}_k)$. Por atualização das estimativas para as comunalidades $\hat{h}_i^2$, ou seja,

$$\tilde{h}_i^2=\sum_{j=1}^{k}\hat{\tau}_{ij}^2$$

Obtêm-se as estimativas de $\mathbf{\Psi}$ por:

$$\tilde{\Psi}=\left(1-\tilde{h}_i^2\right)$$

e se repete este processo (iterativamente) substituindo $\tilde{\mathbf{\Psi}}$ em (13) até que resultados satisfatórios sejam alcançados (CAO, 2010; JOHNSON; WICHERN, 2007).

2.2.4 Interpretações de cargas fatoriais e escores fatoriais

A análise fatorial é uma técnica de interdependência nas quais todas as variáveis são simultaneamente consideradas, pois cada uma relacionada com todas as outras, empregando ainda o conceito da variável estatística, a composição linear de variáveis. Na análise fatorial, as variáveis estatísticas (fatores) são formadas para maximizar seu poder de explicação do conjunto inteiro de variáveis, e não para prever variável (eis) dependente (s).

As cargas fatoriais são as correlações entre as variáveis originais e os fatores. Talvez uma das tarefas mais difíceis na análise fatorial é interpretar os fatores e cargas fatoriais. A

interpretação das cargas fatoriais é relativamente simples à medida que são semelhantes aos coeficientes de regressão na análise de regressão. Esse é um dos pontos principais da análise fatorial, ou seja, quanto maior a carga fatorial maior será a correlação com determinado fator. Um valor negativo indica um impacto inverso no fator.

No entanto, tem-se mais interesse nas implicações das cargas fatoriais, ou seja, nos escores fatoriais. O significado estatístico de cada elemento na matriz de carga fatorial é a de "força" (correlação) de relação de alguma variável observada com alguma variável latente. Suponha algumas variáveis observadas que sejam altamente carregadas em um fator, enquanto as outras têm cargas fatoriais perto de zero em relação a esse fator. Em tal caso, tenta-se interpretar os escores fatoriais correspondentes em termos dessas variáveis observadas.

Muitas vezes, é comum se rotacionar as cargas fatoriais, na tentativa de obter alguma interpretação intuitivamente significativa dos escores fatoriais. Isso, às vezes, é visto como um benefício prático uma vez que oferece uma flexibilidade ao interpretar sem afetar a validade do modelo fatorial. Isso pode ser criticado por sua subjetividade já que a escolha do método de rotação é livre.

Uma estratégia de rotação utilizada é a rotação varimax, que pode resultar em uma "estrutura simples" para um modelo fatorial. Normalmente, cada matriz de carga é dimensionada para um comprimento de forma a alcançar a estabilidade computacional. O objetivo é o de atingir a carga fatorial com matriz de muitos carregamentos, próximas de zero quanto possível. Existem algumas outras técnicas de rotação, como oblíqua e quartimin (CAO, 2010; MARDIA; KENT; BIBBY, 1979; NEUHAUS; WRIGLEY, 1941; THURSTONE, 1938).

2.3. Modelos fator analíticos (FA)

O modelo fator analítico (FA) é baseado na técnica multivariada de análise de fatores (MARDIA; KENT; BIBBY, 1988). No sentido geral, a análise fatorial é usada para modelar a estrutura de covariância entre um conjunto de p variáveis, $y_1,...,y_p$ com o objetivo de explicar as covariâncias em termos de um número menor de fatores hipotéticos. Smith, Cullis e Thompson (2001) usou a abordagem da análise fatorial para fornecer uma estrutura de variância para a matriz de variância genética $\mathbf{G}$.

O modelo é definido em termos de efeitos de genótipos não observados em cada ambiente, como:

$$u_{g_{ij}} = \sum_{t=1}^{k} \tau_{jt} f_{rt} + \delta_{ij} \quad , \qquad (17)$$

em que: $u_{g_{ij}}$ é o efeito aleatório genotípico i $(i = 1,...,m)$ no ambiente j $(j = 1,...,\mathrm{p})$, f_{rt} é o valor (ou escore) do genótipo i no t-ésimo fator hipotético $(t = 1,...,\mathrm{k})$, τ_{jt} é o coeficiente (referente a carga fatorial) para ambiente j e δ_{ij} é o resíduo. Os resíduos δ_{ij} são independentes, com variâncias ψ_j (também conhecidas como variâncias específicaespecíficas no ambiente j). `

O modelo pode ser representado em notação matricial como:

$$\mathbf{u}_{g_{ij}} = (\boldsymbol{\tau}_1 \otimes \mathbf{I}_m)\mathbf{f}_1 + (\boldsymbol{\tau}_2 \otimes \mathbf{I}_m)\mathbf{f}_2 + .., + (\boldsymbol{\tau}_k \otimes \mathbf{I}_m)\mathbf{f}_k + \boldsymbol{\delta},$$

em que que o vetor $(\boldsymbol{\tau}_k \otimes \mathbf{I}_g)\mathbf{f}_k$ é da ordem $\mathrm{mp} \times 1$, $\boldsymbol{\tau}_k$ é o vetor $p \times 1$ das cargas fatoriais $\{\tau_{jt}\}$ é o vetor $\mathbf{f}_k$ $m \times 1$ de escores fatoriais $\{f_{it}\}$, $\boldsymbol{\delta}$ é o $mp \times 1$ vetor dos resíduos (variância específicaespecíficas), $\boldsymbol{\Gamma}$ é a $p \times k$ matriz de cargas, $\{\tau_1,...,\tau_k\}$ e $\mathbf{f}$ é $mk \times 1$, vetor do escores fatoriais, $\left[\mathbf{f}_1^{\mathrm{T}}\mathbf{f}_1^{\mathrm{T}}...\mathbf{f}_k^{\mathrm{T}}\right]^{\mathrm{T}}$.

A distribuição conjunta de $\mathbf{f}$ e $\boldsymbol{\delta}$ é dada por:

$$\begin{pmatrix} \mathbf{f} \\ \boldsymbol{\delta} \end{pmatrix} \sim N\left[\begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} G_f \otimes \mathrm{I}_\mathrm{m} & 0 \\ 0 & \boldsymbol{\Psi} \otimes \mathrm{I}_\mathrm{m} \end{pmatrix}\right]$$

em que: $\boldsymbol{\Psi}$ é a matriz diagonal de variâncias específicas, para cada ambiente, isto é, $\boldsymbol{\Psi} = diag(\psi_1,...,\psi_p)$, os escores fatoriais são assumidas comumente como independentes e com variância unitária, de modo que $G_f \otimes \mathrm{I}_k$. Por conseguinte, a matriz de variâncias e covariâncias para os efeitos genotípicos em cada ambiente é dada pela:

$$\begin{aligned} \mathrm{Var}(u_{g_{ij}}) &= (\boldsymbol{\Gamma} \otimes \mathbf{I}_g)V(\mathbf{f})(\boldsymbol{\Gamma}^{\mathrm{T}} \otimes \mathbf{I}_g) + V(\boldsymbol{\delta}) \\ &= (\boldsymbol{\Gamma} \otimes \mathbf{I}_g)(\mathbf{I}_k \otimes \mathrm{I}_g)(\boldsymbol{\Gamma}^{\mathrm{T}} \otimes \mathbf{I}_g) + (\boldsymbol{\Psi} \otimes \mathbf{I}_g) \\ &= (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}} + \boldsymbol{\Psi}) \otimes \mathbf{I}_g = FA(k) \otimes \mathbf{I}_g \end{aligned}$$

Portanto, o modelo fator de analítico resulta na seguinte forma para $\mathbf{G}$:

$$\mathbf{G} = (\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}} + \boldsymbol{\Psi}) \otimes \mathbf{I}_g = FA(k) \otimes \mathbf{I}_g$$

de modo que as variações genéticas e covariâncias são dados por:

$$\sigma_{g_{jj}} = \sum_{t=1}^{k} \tau_{jt}^2 + \Psi_j$$

$$\sigma_{g_{ij}} = \sum_{t=1}^{k} \tau_{jt}\tau_{it}$$

Em que $\sigma_{g_{jj}}$ é a variância genética do ambiente j e $\sigma_{g_{ij}}$ é a covariância entre o ensaio i e o ensaio j, isto é, $\mathbf{G} = \left\{\sigma_{g_{ij}}\right\}$.

Smith, Cullis e Thompson (2001) mostraram que, quando $k > 1$, $\boldsymbol{\Gamma}$ não é necessariamente única e que $k(k-1)/2$ restrições independentes precisam ser impostas aos elementos de $\boldsymbol{\Gamma}$ para garantir solução única. Os autores ainda definiram $k(k-1)/2$ elementos do triângulo superior de $\boldsymbol{\Gamma}$ serem zero. Isso é $\tau_{jt} = 0$ para $j < t = 2,...k$. Assim, para $k = 2$, o parâmetro τ_{12} é definido como zero. A restrição acima para $\boldsymbol{\Gamma}$ permite uma solução única a ser calculada, mas não tem qualquer significado biológico. Smith, Cullis e Thompson (2001) também sugerem a utilização de uma rotação (com base em componentes principais) das cargas, uma vez estimado como acima, afim de se obter uma interpretação das cargas ambientais e escores genotípicos. Essa rotação assegura que o primeiro fator explica a covariância genética máxima entre ambientes, o segundo fator é responsável pela próxima maior quantidade e é ortogonal em relação ao primeiro fator, e assim por diante.

O modelo de fator analítico para os efeitos genotípicos, dado em (17), é semelhante a um modelo padrão de regressão aleatória, mas com a diferença de que ambas as covariáveis e coeficientes de regressão são desconhecidos e deve ser calculadas a partir dos dados.

Diante das vantagens sugestivas dos modelos FA tem sido aplicado esta estrutura dentro de outras classes de MLB, como AMMI e SREG, para aproximar a estrutura de covariância GE (PIEPHO, 1997, 1998; SMITH; CULLIS; THOMPSON, 2001, 2005; SMITH et al., 2002). E para incorporar informação de parentesco genético como em trabalhos conduzidos por Crossa et al. (2006) e Burgueño et al. (2008). Kelly et al. (2009) e Oakey et al. (2007) que descrevem como modelar covariância GEI e GGEI usando o modelo de FA e como incorporar a matriz aditiva (de parentesco $\mathbf{A}$) e matriz aditivo × aditivo de covariância para o modelo de FA com base em informações de pedigree.

2.4. A relação entre o FA e SREG para a avaliação de GEI

No modelo de FA, o efeito aleatório do genótipo i, no ambiente de ordem j (g_{ij}) é expresso como uma função linear de variáveis latentes $\mathbf{f}_{ik}$ com coeficientes τ_{jk} para $k = (1,2, \ldots t)$, e um resíduo, δ_{ij}, ou seja, $g_{ij} = \mu_j + \sum_{k=1}^{t} f_{ik}\tau_{jk} + \delta_{ij}$ de modo que a j-ésima média de células pode ser escrito como $y_{ij} = g_{ij} + \varepsilon_{ij}$. Se apenas os dois primeiros fatores latentes são retidos, g_{ij} é aproximado por $g_{ij} \approx \tau_1\mathbf{f}_1 + \tau_2\mathbf{f}_2 + \delta_{ij}$. Portanto, existe uma ligação clara entre o SREG2 (modelo SREG com dois eixos retidos para explicar o padrão da GGEI) e modelos FA2 (modelo FA com retenção de duas cargas). Uma conexão similar entre o AMMI e modelos FA foi demonstrada por Smith et al. (2002).

Crossa (2012) nos diz que sob a rotação de componentes principais, as orientações e as projeções dos vetores do FA2 e SREG2 no biplot são os mesmos. Por conseguinte, a propriedade do SREG pelo qual o primeiro componente principal de SREG2 explica a adaptabilidade e o segundo componente principal da SREG2 é devido à variabilidade da GEI, ou seja, explica a estabilidade e deve ser a mesma interpretação dada ao FA2. Deve-se salientar que os valores absolutos dos escores genotípicos e ambientais, sob o FA2 e modelos SREG2 podem não ser necessariamente as mesmas (CROSSA, 2012; STEFANOVA; BUIRCHELL, 2010); as estimativas dos efeitos aleatórios no modelo FA2 são EBLUPs (Melhor preditor linear não viesado), ao passo que as estimativas do modelo SREG2 de efeitos fixos são estimativas de mínimos quadrados, isto é, melhor estimador linear não viesado (EBLUE). Além disso, os erros padrões das funções estimáveis de efeitos fixos sob SREG diferem daquelas de funções preditas de um modelo de efeitos fixos e aleatórios sob FA, e modelos FA são mais flexíveis na manipulação de dados desbalanceados, pois o modelo SREG não lida com dados em falta (BURGUEÑO et al., 2008; CROSSA, 2012). Dessa forma dependendo, de como é feito o ajuste do modelo FA, ele pode ser considerado SERG-misto se o efeito de G for confundido com GEI e AMM-Misto, se o ajuste de G e GEI forem feitos separadamente.

2.5. Relação entre os métodos de análise de componentes principais (PCA) e o modelo fator analítico (FA)

O método de análise de componentes principais (PCA) e modelos fator analítico (FA) proporcionam uma estrutura altamente parcimoniosa para a comparação padrão de matrizes em modelos multi-caracteristicas (*multi-trait*) e, portanto, atraem considerável interesse pelo seu potencial. Ambas as abordagens decompõem a matriz de covariância genética em matrizes de

autovalores e autovetores. Cada vetor singular $\mathbb{u}$, ou seja, componentes principais (PC) forma uma combinação linear das características, enquanto que o correspondente autovalor dá a variância explicada. Os PC são independentes uns dos outros (SCHAEFFER, 1994; TYRISEVÄ et al., 2011).

O objetivo do método PCA é o de detectar todos os componentes necessários para explicar a variação em dados multidimensionais, sem perder qualquer informação importante. O primeiro PC explica a quantidade máxima de variabilidade genética nos dados e todos os PCs, sucessivamente explicam a quantidade da variabilidade máxima restante. Para características altamente correlacionadas, apenas um PC exerce influência sobre a variância genética e PC´s com um efeito insignificante podem ser omitidos sem comprometer a precisão da estimativa. Além disso, a redução de parâmetro resulta numa redução em posição, além de uma redução da dimensão das equações de modelo misto.

O método de FA está relacionado com o método de PC, mas a sua abordagem é diferente. As características estudadas são assumidas ser combinações lineares de algumas variáveis latentes, referidas como fatores comuns. Qualquer variação não explicada por estes é modelada separadamente, ou seja, como características específicas, cabendo correspondentes fatores específicos. Devido ao particionamento da variância em variância comum e específica, o número de fatores necessários para explicar a variabilidade dos dados é normalmente menor do que o número de PC necessários na abordagem PCA (TYRISEVÄ et al., 2011). Além disso, uma vez que os fatores são assumidos como não sendo correlacionados, esparsamente a equação do modelo misto (MME) é obtida como na análise multivariada padrão não estruturada. Contudo, a matriz de covariância resultante é de posto completo se todas as variâncias específicas forem diferentes de zero. Além disso, eixos fatoriais podem ser rotacionados para facilitar a sua interpretação, mas também faz com que seja possível utilizar a parametrização Cholesky (MEYER, 2009; TYRISEVÄ et al., 2011) que aumenta a taxa de convergência dos estimadores de máxima verossimilhança, por exemplo.

2.6. Análise de Fatores sob modelos mistos multiplicativos

Neste tópico iremos mostrar como é feita a extensão do modelo de FA dentro do modelo misto para formulação do modelo fator analítico (FA). O FA é uma forma parcimoniosa usada para aproximar a forma totalmente não estruturada da matriz de covariância genética ($\boldsymbol{\Sigma}$) no modelo de dados MET (KELLY et al., 2007).

Em um modelo FA padrão, um vetor de variáveis aleatórias ($\mathbf{u}$) é descrito como uma combinação linear de um número menor de variáveis aleatórias não observáveis chamados

fatores comuns ($\mathbf{f}$). A equação para i indivíduos (genótipos), quando k fatores comuns são considerados para a modelagem das p variáveis observadas pode ser reescrito a partir de (10) como:

$$\mathbf{u}_i = \boldsymbol{\Gamma}\mathbf{f}_i + \boldsymbol{\delta}_i \tag{18}$$

em que:

$\boldsymbol{\Gamma} = \begin{bmatrix} \tau_{11} & \tau_{12} & \dots & \tau_{1k} \\ \tau_{21} & \tau_{22} & \dots & \tau_{2k} \\ \vdots & \vdots & \dots & \vdots \\ \tau_{m1} & \tau_{m2} & \dots & \tau_{mk} \end{bmatrix}_{m\times k}$ é a matriz ($m\times k$) de cargas fatoriais, $\mathbf{f}_i$ $(k\times 1)$ o vetor dos fatores comuns para cada observação e $\boldsymbol{\delta}_i \left(p\times 1\right)$ o vetor dos fatores específicos para cada observação. Juntando todas as observações a equação (18) pode ser reescrita como:

$$\mathbf{u} = (\boldsymbol{\Gamma}_{\mathrm{p}} \otimes \mathbf{I}_{\mathrm{m}})\mathbf{f} + \boldsymbol{\delta} \tag{19}$$

em que $\mathbf{u} = (u_1^{\mathrm{T}},...,u_m^{\mathrm{T}}) = \begin{pmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_{pi} \end{pmatrix}_{pm\times 1}$, $\mathbf{f} = \left(\mathbf{f}_1^{\mathrm{T}},...,\mathbf{f}_n^{\mathrm{T}}\right) = \begin{pmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \\ f_m \end{pmatrix}_{mk\times 1}$ e

$\boldsymbol{\delta} = \left(\delta_1^{\mathrm{T}},...,\delta_m^{\mathrm{T}}\right) = \begin{pmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_{pi} \end{pmatrix}_{pm\times 1}$.

A equação (18) pode ser vista como um modelo de regressão múltipla multivariada em que o fator aleatório, os escores e a matriz de incidência ($\boldsymbol{\Gamma}$) não são observáveis. Por causa disto, o pressuposto padrão para a identificação do modelo linear é necessário, ou seja, $\boldsymbol{\delta}_i \perp \mathbf{f}_i$ não é suficiente.

Se $\mathbf{H}$ é uma matriz qualquer não singular de ordem apropriada podemos escrever as expressões $\boldsymbol{\Gamma}\mathbf{f} = \boldsymbol{\Gamma}\mathbf{H}\mathbf{H}^{-1}\mathbf{f} = \boldsymbol{\Gamma}^*\mathbf{f}^*$ em que $\boldsymbol{\Gamma}^* = \boldsymbol{\Gamma}\mathbf{H}$ e $\mathbf{f}^* = \mathbf{H}^{-1}\mathbf{f}$. Isto implica que (18) pode ser reescrita como:

$$\mathbf{u}_i = \boldsymbol{\Gamma}^{*}\mathbf{f}_i^{*} + \boldsymbol{\delta}_i \tag{20}$$

de modo que nem $\boldsymbol{\Gamma}^{*}$ e $\mathbf{f}^{*}$ são únicas. No modelo fatorial ortogonal esse problema é resolvido por identificação assumindo que os fatores comuns são mutuamente não correlacionados. No entanto, mesmo com essa hipótese, fatores são determinados só por uma transformação ortonormal. Para verificar isso, considere $\mathbf{T}^{*}$ uma matriz ortonormal tal que $\mathbf{T}^{\mathrm{T}}\mathbf{T} = \mathbf{I}$, então de (18) tem se, $\mathbf{cov(u_i) = \Sigma_u = \Gamma\Gamma^{\mathrm{T}} + \Psi = \Gamma T^{\mathrm{T}} T\Gamma^{\mathrm{T}} + \Psi = \Gamma^{*}\Gamma^{*\mathrm{T}} + \Psi}$,ou seja, obtemos a expressão (20) em que $\boldsymbol{\Psi} = \mathrm{cov}(\boldsymbol{\delta}_i)$ e $\boldsymbol{\Gamma}^{*} = \boldsymbol{\Gamma}\mathbf{T}^{\mathrm{T}}$ (CAMPOS; GIANOLA, 2007; JOHNSON; WICHERN, 2007; MARDIA; KENT; BIBBY, 1979; MEYER, 2009).

Isso significa que, para realizar a identificação, cargas do fator têm de ser rotacionadas numa direção q-dimensional arbitrária. As restrições discutidas acima são arbitrárias e não com base no conhecimento de fundo e em virtude disso, o método é particularmente útil para análise exploratória (CAMPOS; GIANOLA, 2007; JOHNSON; WICHERN, 2007). Além das restrições descritas acima, o método de máxima verossimilhança ou de Inferência Bayesiana exigem que suposições de distribuição de probabilidade sejam assumidas. A distribuição padrão assumida é a Gaussiana (independente e identicamente distribuída), com fatores ortogonais:

$$\begin{pmatrix}\mathbf{f}_i \\ \boldsymbol{\delta}_i\end{pmatrix} \sim \mathbf{N}\left(\begin{pmatrix}\mathbf{0} \\ \mathbf{0}\end{pmatrix}, \begin{pmatrix}\mathbf{I_q} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Psi}\end{pmatrix}\right), \tag{21}$$

sendo $\boldsymbol{\Psi}$, de ordem $p \times p$, é assumido como sendo uma matriz diagonal. Combinando (18) e (21), obtemos a distribuição marginal de $\mathbf{u}_i$:

$$\mathbf{u}_i \sim N\left[0, \boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}} + \boldsymbol{\Psi}\right], \tag{22}$$

Um modelo genético aditivo multivariado padrão para m caracteres (genótipos), medido em cada um vários dos p ambientes pode ser representado pela seguinte expressão:

$$\mathbf{y_i} = \mathbf{X_i}\boldsymbol{\beta} + \mathbf{Z_i}\mathbf{u_i} + \boldsymbol{\varepsilon_i}, \tag{23}$$

em que $\mathbf{y}_i = \left(\mathbf{y}_{i1}, ..., \mathbf{y}_{ip}\right)^{\mathrm{T}}$ é o ($p \times 1$) vetor de medidas fenotípicas avaliadas no genótipo i (i = 1, ..., m); $\boldsymbol{\beta}$ e $\mathbf{u}$ são vetores desconhecidos de coeficientes de regressão e de efeitos genéticos aditivos, respectivamente. $\mathbf{X}_i$ e $\mathbf{Z}_i$ são as matrizes de incidência conhecidas de ordem

apropriada, e $\boldsymbol{\varepsilon_i}$ é o $(p \times 1)$ vetor de resíduos do modelo. Juntando todas as observações, a equação o conjunto de dados (em notação matricial) é,

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \tag{24}$$

em que $\mathbf{y} = \left(\mathbf{y}_1^{\mathrm{T}}, ..., \mathbf{y}_p^{\mathrm{T}}\right)^{\mathrm{T}}$, $\mathbf{X} = [\mathbf{X}_1^{\mathrm{T}}, ..., \mathbf{X}_p^{\mathrm{T}}]^{\mathrm{T}}$, $\mathbf{Z} = diag\{\mathbf{Z}_i\}$, $\mathbf{u_i} = (\mathbf{u}_1^{\mathrm{T}}, ..., \mathbf{u}_p^{\mathrm{T}})^{\mathrm{T}}$ e $\boldsymbol{\varepsilon} = \left(\boldsymbol{\varepsilon}_1^{\mathrm{T}}, ..., \boldsymbol{\varepsilon}_p^{\mathrm{T}}\right)^{\mathrm{T}}$.

A suposição de distribuição padrão em genética quantitativa é:

$$\begin{pmatrix} \boldsymbol{\varepsilon} \\ \mathbf{u} \end{pmatrix} \sim \mathbf{N}\left(\mathbf{0}, \begin{pmatrix} \mathbf{I}_m \otimes \mathbf{R} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_m \otimes \mathbf{G} \end{pmatrix}\right), \tag{25}$$

sendo $\mathbf{R}$ e $\mathbf{G}$ matrizes de variância e covariâncias ($p \times p$)de resíduos do modelo e dos efeitos genéticos aditivos, respectivamente, e $\mathbf{I}$ é a matriz de Identidade ($m \times m$). Assumindo que (18) é válida para o vetor de efeitos genéticos aditivos em (24) tem se:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\left(\boldsymbol{\Gamma}_{\mathrm{p}} \otimes \mathbf{I}_{\mathrm{m}}\right)\mathbf{f} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}\ , \tag{26}$$

em que $\boldsymbol{\Gamma}$ é, como definido anteriormente, e **f** e **δ** são interpretados como vetores comum e vetores de efeitos genéticos aditivos específicos, respectivamente.

Combinando os pressupostos do modelo FA ortogonal descrito acima com os do modelo genético aditivo leva à distribuição conjunta:

$$\begin{pmatrix} \boldsymbol{\varepsilon} \\ \mathbf{f} \\ \boldsymbol{\delta} \end{pmatrix} \sim \mathbf{N}\left(\mathbf{0}, \begin{pmatrix} \mathbf{I}_{\mathrm{m}} \otimes \mathbf{R} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}_{\mathrm{m}} \otimes \boldsymbol{\Psi} \end{pmatrix}\right), \tag{27}$$

em que $\boldsymbol{\Psi}$ (p× p) representa a matriz de variâncias e covariâncias de efeitos genéticos aditivos específicos. Nota-se que, em (19), ao contrário da norma do modelo FA, ou seja, (25), com diferentes níveis de fatores comuns e específicos estes podem estar correlacionados devido às relações genéticas. Com estas premissas, a distribuição condicional dos dados, dado **β**, **u** e **R** é dada por:

$$\mathbf{y} \mid \mathbf{u}, \boldsymbol{\beta}, \mathbf{R}_0 \sim \mathbf{N}\left[\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u}, \mathbf{I} \otimes \mathbf{R}\right]. \tag{28}$$

Alternativamente, utilizando (26), pode-se escrever:

$$\mathbf{y} \mid \mathbf{u}, \boldsymbol{\beta}, \mathbf{R} = \mathbf{y} \mid \mathbf{f}, \boldsymbol{\delta}, \boldsymbol{\Gamma}, \boldsymbol{\beta}, \mathbf{R} \sim N\left[\mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\left(\boldsymbol{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}}\right)\mathbf{f} + \mathbf{Z}\boldsymbol{\delta}, \mathbf{I} \otimes \mathbf{R}\right]$$

Os parâmetros são estimados usando o procedimento REML/BLUP, em que os componentes de variância são estimados pela máxima verossimilhança restrita (REML), e os valores genotípicos preditos pelo melhor preditor linear não viciada (BLUP).

Dada a matriz de variância genética ($\boldsymbol{\Sigma}$) e os BLUPs dos efeitos de genótipos e assumindo $\boldsymbol{\Sigma}$ **(Σ=G)** como uma estrutura de FA, ou seja, ($\boldsymbol{\Gamma\Gamma}^{\mathrm{T}} + \boldsymbol{\Psi}$), e ainda observando que os BLUPs podem ser representados por fatores comuns na forma ($\mathbf{u} = \mathbf{Lf} + \boldsymbol{\delta}$, com $\mathbf{L} = \left(\boldsymbol{\Gamma}_{\mathrm{p}} \otimes \mathbf{I}_{\mathrm{m}}\right)$), um modelo equivalente a (26) pode ser reescrito para facilitar as operações algébricas, o qual é dado por:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}[\mathbf{Lf} + \boldsymbol{\delta}] + \boldsymbol{\varepsilon} \ . \qquad (28a)$$

Assim, a solução matriz das equações de modelos mistos reparametrizadas ($\mathbf{W} = \mathbf{ZL}$) pode ser dada por:

$$\begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} & \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} + \mathbf{I} & \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} & \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} + \boldsymbol{\Psi}^{-1} \otimes \mathbf{I} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{b}} \\ \hat{\mathbf{f}} \\ \hat{\boldsymbol{\delta}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \end{bmatrix}. \qquad (29)$$

Resolvendo (29) em relação $\hat{\mathbf{b}}$, $\hat{\mathbf{f}}$ e $\hat{\boldsymbol{\delta}}$ tem-se:

$$\begin{bmatrix} \hat{\boldsymbol{\beta}} \\ \hat{\mathbf{f}} \\ \hat{\boldsymbol{\delta}} \end{bmatrix} = \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} & \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} + \mathbf{I} & \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} \\ \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{X} & \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{W} & \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{Z} + \boldsymbol{\Psi}^{-1} \otimes \mathbf{I} \end{bmatrix}^{-1} \begin{bmatrix} \mathbf{X}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{W}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \\ \mathbf{Z}^{\mathrm{T}}\mathbf{R}^{-1}\mathbf{y} \end{bmatrix} \qquad (30)$$

Assumindo a matriz de covariância genética modelada pela estrutura FA, $\left(\boldsymbol{\Sigma}=\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}+\boldsymbol{\Psi}\right)$ a solução dos efeitos fixos e aleatórios os estimadores dos escores fatoriais (EBLUP´s) e das variâncias específicas $\hat{\mathbf{f}}$ e $\hat{\boldsymbol{\delta}}$, segundo Meyer (2009) são obtidos de forma similar a solução da equação dos modelos mistos.

A vantagem de usar este procedimento está na garantia da convergência do modelo dentro do espaço do parâmetro evitando a ocorrência dos casos Heywood.

2.7. Estruturas de covariâncias comumente utilizadas para modelar a GEI

Modelos lineares mistos consistem em duas partes: um dos efeitos fixos, que representa a esperança dos valores observados e a outra, a aleatória que representa a covariância das observações. Da específicação funcional da estrutura de covariância para o modelo misto é feito ...por meio ... de **G** e **R,** segundo **o** modelo (23).

Existem diversas estruturas de covariâncias residuais que podem ser selecionadas para representar a variação de medidas repetidas em uma mesma unidade experimental. Dentre os diferentes tipos de estruturas de covariâncias estudadas podem ser citadas: AR(1) - autorregressiva de primeira ordem; VC - componentes de variância; CS - simetria composta; UN - não-estruturada; ARH(1) - autorregressiva heterogênea de primeira ordem; FA(1) - fator analítico de primeira ordem; HF - Huynh-Feldt; TOEP - Toeplitz; TOEPH - Toeplitz heterogênea. Na estrutura CS, todas as covariâncias são iguais e, na VC, as covariâncias são nulas.

Dentre esta estrutura de covariância, a mais simples assume que todas as observações são independentes e não existe correlação (covariância) entre qualquer par de observações, mesmo entre as medidas repetidas sobre o mesmo ambiente. A estrutura simples tem uma variância σ^2 igual na diagonal principal e os outros elementos da matriz tem valores iguais ao zero, e é representada como:

$$\begin{bmatrix} \sigma^2 & 0 & 0 & 0 \\ 0 & \sigma^2 & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma^2 \end{bmatrix}.$$

Esta estrutura é considerada simples porque apenas uma única estimativa do parâmetro é obrigatória. Esta é a estrutura de covariância assumida na análise de variância padrão do modelo fixo. Uma generalização desta é a estrutura diagonal, onde ainda são independentes, mas cada um pode ter variância diferente em cada ambiente:

$$\begin{bmatrix} \sigma_{11} & 0 & 0 & 0 \\ 0 & \sigma_{22} & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_{ii} \end{bmatrix}$$

2.7.1Estrutura de covariância simetria composta (CS)

A simetria composta refere-se a variâncias iguais na diagonal ($\sigma^2+\sigma_1$) e covariâncias iguais (σ_{11}) fora das diagonais. Essa estrutura mais simples é a estrutura de covariância correlacionada para dados de medidas repetidas ou MET, uma vez que assume uma correlação constante entre as observações, independentemente da distância entre pontos de tempo ou espaço. Nesta estrutura temos igualdade de variâncias e covariâncias, ou seja, covariâncias constantes entre quaisquer observações de uma mesma unidade devido a erros independentes. A estrutura CS requer estimação de dois parâmetros de variâncias, entre ambientes (σ_{11}) e a intra-ambientes (σ^2) e matriz de covariância sendo expressa como:

$$\begin{bmatrix} \sigma^2+\sigma_{11} & \sigma_{11} & \sigma_{11} & \sigma_{11} \\ \sigma_{11} & \sigma^2+\sigma_{11} & \sigma_{11} & \sigma_{11} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{11} & \sigma_{11} & \sigma_{11} & \sigma^2+\sigma_{11} \end{bmatrix}.$$

A estrutura de covariância CS só é apropriada quando a condição denominada Huynh-Feldt for cumprida, que é igual correlação entre as medidas sobre o mesmo ambiente (HUYNH; FELDT, 1970, 1976). Se a condição HF não for cumprida, outras estruturas de covariâncias devem ser testadas.

2.7.2 Estrutura covariância AR (1)

A estrutura de covariância autorregressiva de primeira ordem assume a correlação entre as medidas adjacentes é (ρ), independentemente da ordem dos pares adjacentes, como 1º e 2º, 2º e 3º, e assim por diante. Ela também pressupõe que a correlação para qualquer par de observações que são medidos n unidades para além de ter uma correlação ρ^n. A correlação entre as observações é uma função da distância no tempo. Além disso, ela assume variâncias iguais σ^2 na diagonal. A variância e a correlação correspondente fora das diagonais e matriz de covariância são expressas como:

$$\sigma^2\begin{bmatrix} 1 & \rho & \rho^2 & \rho^3 \\ \vdots & 1 & \rho & \rho^2 \\ \vdots & \vdots & 1 & \rho \\ \vdots & \vdots & \vdots & 1 \end{bmatrix}$$

Na estrutura AR (1), as correlações diminuem com o aumento das distâncias entre pares de observações e as covariâncias correspondentes diminuem (covariância entre duas observações decresce a medida em que aumenta o intervalo de tempo entre elas). A estrutura AR (1) requer tempos igualmente espaçados, e devem ser corretamente ordenados e a estrutura precisa de apenas duas estimativas dos parâmetros. Se tempos desigualmente espaçados estiverem presentes, devem-se considerar outras estruturas de covariância, como potência e esféricas.

2.7.3 Estrutura Covariância toeplitz

a) TOEP: Toeplitz

Utilizada para dados de séries temporais igualmente espaçados e correlação arbitrária para cada defasagem. A estrutura TOEP é semelhante ao AR (1) em que todas as medições ao lado, uma da outra, têm a mesma correlação. Duas medições separadas têm a mesma correlação diferente da primeira, três medições têm a mesma correlação diferente das duas primeiras etc. (não se uso vírgula antes de (etc.). No entanto, as correlações não têm necessariamente o mesmo padrão como no AR (1). Tecnicamente, a AR (1) é um caso especial da Toeplitz e expressa como:

$$\begin{bmatrix} \sigma_{11} & \sigma_2 & \sigma_3 & \sigma_4 \\ & \sigma_{11} & \sigma_2 & \sigma_3 \\ & & \sigma_{11} & \sigma_2 \\ & & & \sigma_{11} \end{bmatrix}$$

b) TOEPH: Toeplitz heterogênea

Utilizada para dados de séries temporais igualmente espaçados, com parâmetros de variâncias diferentes para cada elemento da diagonal, sendo os elementos fora da diagonal principal funções de variâncias e do k-ésimo parâmetro de autocorrelação (ıρk1< ı) para cada defasagem q-1, e zeros para as últimas defasagens, e é expressa como:

$$\begin{bmatrix} \sigma_{11} & \rho\sigma_1\sigma_2 & \dots & 0 \\ \rho\sigma_2\sigma_1 & \sigma_{22} & \vdots & \rho\sigma_i\sigma_2 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \sigma_{ii} \end{bmatrix}$$

2.7.4 HF: Huynh-Feldt

Essa estrutura é similar à simetria composta heterogênea, que tem o mesmo número de parâmetros e heterogeneidade ao longo da diagonal principal. Entretanto, a construção dos elementos fora da diagonal é feita tomando-se a média aritmética entre as variâncias e subtraindo ξ, em que ξ é a diferença entre a média das variâncias e a média das covariâncias. Essa estrutura é expressa como:

$$\begin{bmatrix} \sigma_{11} & \frac{\sigma_{ii}+\sigma_{22}}{2}-\xi & \cdots & \frac{\sigma_{11}+\sigma_{ii}}{2}-\xi \\ \frac{\sigma_{22}+\sigma_{11}}{2}-\xi & \sigma_{22} & \vdots & \frac{\sigma_{22}+\sigma_{ii}}{2}-\xi \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\sigma_{ii}+\sigma_{11}}{2}-\xi & \frac{\sigma_{ii}+\sigma_{22}}{2}-\xi & \cdots & \sigma_{ii} \end{bmatrix}.$$

2.7.5 Estrutura de covariância não estruturada (UN)

A estrutura de covariância não estruturada permite variâncias desiguais ao longo dos ambientes e covariâncias desiguais para cada combinação de ambientes. Específica uma matriz completamente geral, parametrizada diretamente em termos de variâncias e covariâncias. As variâncias são restritas a valores não negativos e as covariâncias não têm restrições.

Esta é a mais complexa estrutura e $p\left(p-1\right)/2$ parâmetros devem ser estimados em que p é o número de ambientes é expressa como:

$$\begin{bmatrix} \sigma_{11} & \sigma_{12} & \sigma_{13} & \sigma_{1i} \\ \sigma_{21} & \sigma_{22} & \sigma_{23} & \sigma_{i2} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{i1} & \sigma_{i2} & \sigma_{i3} & \sigma_{ii} \end{bmatrix}$$

Devido à complexidade da estrutura UN, é geralmente mais difícil de ajustar, mesmo com um conjunto menor de dados ainda é possível se observar falhas na convergência. Como a estrutura da UN estima muitos parâmetros pode-se perder uma quantidade considerável de informações nos dados e resultar em testes menos poderosos.

2.7.6 Estrutura covariância FA

Especifica uma estrutura com k fatores (JENNRICH; SCHLUCHTER, 1986). Essa estrutura é da forma $\mathbf{\Gamma\Gamma}^{\mathrm{T}} + \mathbf{\Psi}$, onde $\mathbf{\Gamma}$ é uma matriz retangular (p×k) e $\mathbf{\Psi}$ é uma matriz diagonal ($p \times p$) de variâncias únicas. Quando $k > 1$, os elementos no canto superior direito (elementos na i-ésima linha e j-ésima coluna com j > i) de $\mathbf{\Gamma}$ são um conjunto de zeros. É dada por:

$$\begin{bmatrix} \delta_1^2 + \sigma_1^2 & \delta_1\delta_2 & \cdots & \delta_1\delta_i \\ \delta_2\delta_1 & \delta_2^2 + \sigma_2^2 & \cdots & \delta_1\delta_i \\ \vdots & \vdots & \ddots & \vdots \\ \delta_i\delta_1 & \cdots & \cdots & \delta_i^2 + \sigma_i^2 \end{bmatrix}$$

2.7.6.1 FA1(1): Fator analítico de primeira ordem

É similar à estrutura fator analítico exceto que todos os elementos σ_1^2 da diagonal devem ser iguais. Essa é usada em modelo misto se baseando no modelo de Finlay e Wilkinson (1963). Esta estrutura é designada como: FA1(1) (igual estutura fator analítica diagonal de primeira ordem) em modelos mistos, e é representada como:

$$\begin{bmatrix} \delta_1^2 + \sigma_1^2 & \delta_1\delta_2 & \cdots & \delta_1\delta_i \\ \delta_2\delta_1 & \delta_2^2 + \sigma_1^2 & \cdots & \delta_1\delta_i \\ \vdots & \vdots & \ddots & \vdots \\ \delta_i\delta_1 & \cdots & \cdots & \delta_i^2 + \sigma_1^2 \end{bmatrix}.$$

Ainda usando o modelo de Finlay–Wilkinson, Piepho (1997) propôs ajustar no contexto de modelos mistos a estrutura FA. O FA (1) corresponde uma forma específica de um modelo AMMI (GAUCH, 1988), e são denotados como uma estrutura AMMI. A forma da matriz em um modelo misto é o seguinte:

$$\begin{bmatrix} \delta_1^2 + \sigma_1^2 & \delta_1\delta_2 & \cdots & \delta_1\delta_i \\ \delta_2\delta_1 & \delta_2^2 + \sigma_1^2 & \cdots & \delta_1\delta_i \\ \vdots & \vdots & \ddots & \vdots \\ \delta_i\delta_1 & \cdots & \cdots & \delta_i^2 + \sigma_1^2 \end{bmatrix} + \mathbf{J}\sigma_1^2.$$

O FA(1) é flexível pois permite aos pesquisadores escolher quantos fatores devem incluir no modelo. A escolha de muitos fatores permite maior flexibilidade no ajuste do modelo, mas pode reduzir a parcimônia modelo.

Usando essa estrutura de covariância vários pesquisadores demonstraram que a mesma pode ser combinada com informação de pedigree para melhorar o ajuste do modelo, tal como medido por critérios de informação (CROSSA et al., 2006; KELLY et al., 2009; OAKEY et al., 2007). Embora esses pesquisadores tenham usado um conjuto limitado de dados reais MET, com dados de simulação mostraram que o modelo FA com ou sem parentesco (se díponivel) é o modelo mais eficaz para análise de dados MET.

Segundo Smith et al. (2015) num esforço para melhorar as taxas de convergência, os modelos são ajustados sequencialmente da estrutura mais baixa a mais complexa. As estimativas dos parâmetros de modelos mais simples devem ser usadas como os valores de partida da estrutura seguinte mais complexa para a qual a estrutura simples era um caso específico. Se um modelo não converge, a seguinte estrutura mais complexa não deve ser ajustada.

2.7.6.2 Critérios para a escolha da estrutura de matriz de variância e covariâncias

A abordagem de modelos mistos para a análise de dados permite que os usuários escolham a estrutura mais adequada para modelar as covariâncias dos seus dados. Ao invés de usarem testes estatísticos univariados ou multivariados para analisar os efeitos, ou testes que assumem uma forma particular para a estrutura de covariância, a abordagem de modelo misto permite que os dados determinem a estrutura adequada (KESELMAN et al., 1998). O benefício potencial dessa abordagem é que ela permite a um pesquisador modelar a estrutura de covariância dos dados, ao invés de assumir um determinado tipo de estrutura, como é o caso com as estatísticas de teste univariada e multivariada tradicional (LITTELL; PENDERGAST; NATARAJAN, 2000; WOLFINGER, 1993). Parcimoniosamente, modelar a estrutura de covariância dos dados deve resultar em estimativas mais eficientes dos parâmetros de efeitos fixos do modelo e, consequentemente, testes mais poderosos dos efeitos aleatórios (KESELMAN et al., 1998). A abordagem de modelo misto permite que os usuários escolham várias estruturas de covariância com os dados. E Essas escolhas dependem dos critérios de informação (LITTELL; PENDERGAST; NATARAJAN, 2000).

Alguns dos critérios para escolher a estrutura de covariância mais adequada para realização das análises são: os valores das funções de verossimilhança restrita -2RLL (“-2 Res

Log Likelihood"); AIC ("Akaike's Information Criterion"); AICC ("Consistent Akaike's Information Criterion"), BIC ("Bayesian Information Criterion") e QIC ("quasi-information criterion"). Além disso, deve ser verificada a violação do critério de esfericidade da matriz pelo teste esfericidade de "Mauchly". Segundo Xavier (2000), o teste de esfericidade de "Mauchly" apresenta a estatística para avaliar a condição de esfericidade. Essa verifica se uma população normal multivariada apresenta variâncias iguais e as correlações nulas. Caso uma população apresente essa simetria, será chamada de "esférica". Valores de -2RLL, AIC, AICC e BIC mais próximos de zero indicam melhores ajustes. O teste de máxima verossimilhança, além de comparar os modelos aninhados, tende a favorecer os modelos que possuem o maior número de parâmetros. Os critérios AIC, AICC e BIC permitem a comparação de modelos não aninhados e penalizam os modelos com maior número de parâmetros (WOLFINGER, 1993).

Barnett et al. (2010) usando dados de ecologia para selecionar a melhor estrutura de covariância usaram três critérios diferentes (dentre os disponíveis) e os resultados mostraram que o DIC era uma melhor estatística dentre todas para fazer esta escolha, embora tenha sido superado pelo AIC quando a verdadeira estrutura de covariância fosse a independente (Diagonal Principal "Bartded" (UN(1)). Verificaram ainda que para selecionar a covariância ideal quando esses modelos fossem encaixando seria a utilização de qualquer modelo de covariância padrão e o AIC, ou uma abordagem bayesiana e o DIC, não recomendando o uso do QIC, pois esta não penaliza covariâncias suficientemente complexas, e por isso muitas vezes erroneamente seleciona modelos mais complexos. A quasi-verossimilhança usada pelo QIC não é uma boa estatística para detectar as diferenças entre os modelos ajustados.

2.8. Análise bayesiana

Métodos bayesianos vêm aumentando a sua popularidade nas ciências como meio de inferência probabilística (COSTA, 2004; MALAKOFF, 1999). Dentre suas vantagens encontram-se a habilidade de incluir informação prévia (a priori), a facilidade da incorporação desta num contexto formal de decisão, o tratamento explícito da incerteza e a capacidade de assimilar novas informações em contextos adaptativos (COSTA, 2004). Além disso, a inferência bayesiana é atraente porque promove uma interpretação de "senso comum" de conclusões estatísticas. A saída da análise é uma distribuição a posteriori, que fornece instantaneamente a capacidade de estimar intervalos de credibilidade para um parâmetro desconhecido ou para calcular a probabilidade de um evento de interesse. Esta quantificação

direta de incerteza leva à capacidade de ajustar modelos complicados com muitos parâmetros e especificações de vários níveis de probabilidade (GELMAN et al., 2013).

2.8.1. Inferência bayesiana

Inferência estatística refere-se à obtenção de conclusões sobre os parâmetros (quantidades não observadas θ) a partir de dados observados (*y*). A inferência bayesiana aborda o problema definindo a probabilidade de uma forma subjetiva, como uma medida da plausibilidade de uma proposição, condicional no conhecimento do observador. A incerteza em relação a θ pode assumir diferentes graus, os quais se representam ...por meio ... de modelos probabilísticos para θ. Portanto, tanto as quantidades observáveis, quanto os parâmetros do modelo estatístico são considerados quantidades aleatórias. Esta última característica constitui uma diferença fundamental da abordagem bayesiana em relação à clássica, que considera o parâmetro como uma quantidade fixa e desconhecida, à qual nós aproximamos no processo de inferência (BERNARDO; SMITH, 1994; TANNER, 1993).

Do ponto de vista prático, a modelagem bayesiana inicia com a especificação de um modelo probabilístico completo, ...por meio ... da distribuição conjunta das quantidades observáveis e não observáveis do problema (BERNARDO, 2005b; COSTA, 2004; GELMAN et al., 2013). A informação disponível sobre θ, resumida na densidade de probabilidade $p(\theta)$, é aumentada com a observação de uma quantidade aleatória *y* que se relaciona com θ. O teorema de Bayes fornece a regra de atualização desta informação.

2.8.2. Teorema de Bayes

Este teorema estabelece uma relação entre probabilidades condicionais, dado por:

$$P(B|A,C) = \frac{P(A|B,C) \cdot P(B|C)}{P(A|C)} \quad \textbf{(31)}$$

e pode ser interpretado como um modelo do processo de aprendizado sobre θ:

$$p(\theta|y,M) = \frac{p(y|\theta,M) \cdot p(\theta|M)}{p(y|M)} \quad \textbf{(32)}$$

onde:

$p(y|\theta,M)$: Verossimilhança ou distribuição de probabilidade dos dados *y* condicionais nos parâmetros θ e nas hipóteses do modelo *M*, que fornece a plausibilidade de cada um dos possíveis valores do parâmetro θ;

$p(\theta|M)$: Distribuição *a priori*, que expressa o conhecimento sobre os parâmetros antes de examinar os dados;

$p(y|M)$: Evidencia, ou constante de normalização;

$p(\theta|y,M)$: Distribuição *a posteriori* que expressa o conhecimento sobre o parâmetro apôs examinar os dados.

A efeito de simplificar a notação, o condicionamento em relação as hipóteses do modelo *M* será omitido.

2.8.3 Evidência

O denominador da expressão (312) é a soma sobre todos os valores possíveis de θ. $p(y)=\sum_{\theta} p(\theta)\cdot p(y|\theta)$, ou $p(y)=\int_{\theta} p(\theta)\cdot p(y|\theta)\cdot d\theta$ no caso de θ contínuo. Devido a que este fator $p(y)$ não depende de θ, para um conjunto de dados observados específicos pode ser considerado uma constante, de forma que se pode reescrever (31.2) como:

$$p(\theta|y)\propto p(\theta)\cdot p(y|\theta) \tag{33}$$

O lado direito da expressão (31.3) é a densidade *a posteriori* não normalizada ou *kernel* da função densidade de probabilidade *a posteriori* condicional de θ em *y*. Estas expressões resumem a essência da inferência bayesiana. A tarefa fundamental de qualquer aplicação específica é desenvolver o modelo $p(\theta, y)$ e realizar os cálculos necessários para representar o conhecimento *a posteriori* $p(\theta|y)$ de forma apropriada.

2.8.4 Estimação bayesiana e regiões credibilidade

Para descrever o conteúdo inferencial da distribuição a posteriori $p(\theta|\mathbf{y})$ da quantidade de interesse é sempre conveniente citar regiões $\mathbf{R}\subset\mathbf{\Theta}$ de determinada probabilidade (a posteriori sob $p(\theta|\mathbf{y})$).

Qualquer subconjunto do espaço de parâmetros $\mathbf{R}_q\subset\mathbf{\Theta}$ tal que

$$\int_{\mathbf{R}_q} p(\theta \mid \mathbf{y}) d\theta = q, 0 < q < 1 , \qquad (34)$$

de modo que, dado os dados $\mathbf{X}$, o verdadeiro valor de $\boldsymbol{\theta}$ pertence a $\mathbf{R}_q$ com probabilidade q ,é dito ser uma região (a posteriori) de credibilidade-q de $\boldsymbol{\theta}$. Regiões de credibilidade são coerentes sob reparametrização;

Assim, para qualquer região de credibilidade-q $\mathbf{R}_q$ de $\boldsymbol{\theta}$ uma transformação um-para-um $\phi = \phi(\boldsymbol{\theta}), \phi(R_q)$ é uma região e credibilidade-q de ϕ. No entanto, para qualquer q existem geralmente infinitamente muitas regiões de credibilidade.

Às vezes, regiões de credibilidade são selecionadas para terem um tamanho mínimo (comprimento, área, volume), o que resulta em regiões de maior densidade de probabilidade a posteriori (HPD), onde todos os pontos da região têm maior densidade de probabilidade do que todos os pontos fora.

Portanto, em análises estatísticas, é preciso especificar os métodos de estimação, como construir regiões de credibilidade e como fazer avaliações das hipóteses. No contexto Bayesiano, todas as probabilidades relativas aos parâmetros de interesse são baseadas na posteriori. O princípio da probabilidade condicional afirma que a distribuição a posteriori $p(\theta \mid \mathbf{y})$ resume o estado atual do conhecimento sobre θ. Supondo que estamos interessados em fazer inferência sobre uma característica de interesse, ou seja, $\tau = \mathrm{T}(\theta)$ onde $\mathrm{T}: \Theta \to T$. Quaisquer declarações sobre probabilidade τ deve ser condicional em $\mathbf{X}$ e inferências sobre τ são determinados da mesma forma como inferências sobre θ. Em geral, é preciso determinar a distribuição a posteriori marginal de τ .

A medida de probabilidade posteriori de τ é dada por:

$$\pi_T(. \mid X) = \left(\mathrm{T}^{-1}(.) \mid \mathbf{X}\right) \qquad (35)$$

Às vezes, é difícil obter a forma fechada para a densidade da distribuição a posteriori de τ. Em tal situação, podemos utilizar técnicas computacionais para aproximar as experanças dos parâmetros de interesse.

Um procedimento de estimação comum é a máxima a posteriori (MAP). Para isso, é necessário obter a moda da densidade a posteriori de τ , ou seja,

$$\hat{\tau} = \arg\max_{\tau} \pi_{\mathrm{T}}(. \mid \mathbf{X}).$$

Note que a estimativa MAP coincide com a estimativa de máxima verossimilhança quando a priori é uniforme. Também pode ser mostrado pela regra de Bayes ao usar a função de perda $0-1$.

Alternativamente, a média a posteriori, se existir, é utilizada, ou seja,

$$\hat{\tau} = E\left(\mathrm{T}(\theta) \mid \mathbf{X}\right).$$

Esta é uma regra de Bayes (***Bayes rule***), quando a função perda quadrática é usada.

Nós também podemos resumir a informação na distribuição a posteriori por meio de regiões de maior densidade a posteriori (regiões a posterioris HPD). Uma região $\gamma - HPD$ é definida pela $D_{\gamma}(\mathbf{X}) = \left\{\tau : \pi_{\mathrm{T}}(\tau \mid \mathbf{X}) \geq c\right\}$, em que c é o menor valor de modo a que tal região tenha menor volume possível entre todos os subgrupos contendo γ da probabilidade a posteriori quando o espaço de parâmetros é um subconjunto de $\Re^k$. Observa-se que quando a moda de $\hat{\tau}$ é única, então $D_{\gamma}(\mathbf{X}) \downarrow \{\hat{\tau}\}$ como $\gamma \downarrow 0$.

Um defeito principal das inferências baseadas em HPD, tais como a moda, é que elas não são invariantes sob reparameterizações suaves (BERNARDO, 2005a; CAO, 2010). Como exemplo considere a imagem $\phi(R_q)$ de uma $\mathbf{R}_q$ (região de credibilidade-q) HPD que será uma região de credibilidade-q para ϕ, mas geralmente não será HPD; na verdade, não há nenhuma razão para restringir a atenção para regiões de credibilidade HPD. Em uma dimensão, quantiles a posterioris são muitas vezes utilizados para obter regiões de credibilidade. Assim, se $\theta_q = \theta_q(x)$ é o quantil $100q\%$ da posteriori de θ, então

$$R_q^l = \left\{\theta; \theta \leq \theta_q\right\}$$

é uma região de credibilidade-q unilateral, geralmente única, e é coerente sob reparametrização. Probabilidades de regiões de credibilidade-q centradas da forma: $R_q^c = \left\{\theta; \theta_{(1-q)/2} \leq \theta_q \leq \theta_{(1+q)/2}\right\}$ são mais fáceis de calcular, e são frequentemente preferidas em detrimento de regiões HPD (BERNARDO, 2005a; CAO, 2010; ROWE, 2003).

No entanto, regiões de credibilidade centradas são apenas muito atraentes quando a densidade a posteriori tem uma moda singular interior (*posterior density has a unique interior mode*), e tem uma limitação fundamental: eles não são exclusivamente definidos em problemas com mais de uma dimensão (BERNARDO, 2005a).

Para as funções de perda razoáveis, tipicamente uma única região de credibilidade pode ser selecionada como uma *região de menor perda a posteriori* (LPL- *lowest posterior loss region*), onde todos os pontos da região têm menor perda esperada (a posteriori) que todos de pontos fora.

2.9.1 Análise de fatores bayesiana (AFB)

O desenvolvimento da técnica de análise de fatores bayesiana (AFB) pode ser considerado "recente", pois seus estudos ganharam "força" somente a partir dos anos 2000, apesar da técnica ter sido apresentada nos inícios da década 70. Os estudos da AFB podem ser divididos em dois estágios: os que consideram a técnica Pré-MCMC (antes da década 90) e pós-MCMC (depois da década 90). Os estudos pré-MCMC são atribuidos a Geweke e Zhou (1996), Hyashi e Sen (2002), Hyashi e Yung (1999), Lee (1994), Press (1972) e Press e Shigemasu (1989). As abordagens pós-MCMC, por outro lado, podem ser encontradas em Cao (2010), Geweke e Zhou (1996), Gosh e Dunson (2009), Lopes e Carvalho (2007) e Rowe (2002, 2003), que apresentam diversas propostas para o estudo da AFB, apresentando metodologias para atribuição de densidades a prioris e obtenção das posterioris condicionais. Nesta seção será dado destaque a revisão sobre AFB seguindo a metodologia apresentada por Cao (2010) e Rowe (2003).

Em AFB é comum se atribuírem densidades (distribuições) a priori normais para cargas fatoriais e densidades gamas inversas para variâncias residuais, devido à sua forma conjugada condicionalmente que facilita a obtenção de condicionais completas de forma fechada com amostragem direta. No entanto, estas prioris requerem elicitação de muitos hiperparâmetros e tendem a resultar em algoritmos para o amostrador de Gibbs (GS) mal comportados.

Nesta tese iremos apresentar duas abordagens para a AFB. Como em Cao (2010) e Rowe (2003), estas abordagens referem-se primeiro a seleção do modelo (model selection approach) e a segunda a concentração da posteriori (posterior concentration approach). A primeira abordagem requer que seja atribuída priori sobre $(\mu, \mathbf{\Gamma}_q, \mathbf{\Psi})$ assim como em $q \in \{0,...,k\}$. A segunda, por sua vez,requer priori sobre $(\mu, \mathbf{\Sigma})$, mais detalhes sobre essas abordagens podem ser encontrados em Lee (2007). A seguir serão descritas as principais prioris assumidas para

estas abordagens, bem como os aspectos importantes para análises a posteriori (posterior analysis).

2.9.2 Abordagem de seleção de modelos

Nessa abordagem é preciso atribuir $p+1$ probabilidades a priori para que o modelo com i fatores comuns seja correto, ou seja, $\pi\left(M_i\right)$ para $i=0,...,p$ em que:

$$M_i=\left\{P_{\{\mu,\Sigma_i\}},\mu\in\Re^1,\Sigma_i=\mathbf{\Gamma}_i\mathbf{\Gamma}_i^{\mathrm{T}}+\mathbf{\Psi}\right\} \tag{36}$$

É necessário ainda especificar densidades a priori para $(\mu,\mathbf{\Gamma}_i,\mathbf{\Psi})$. Note que quando $\mathbf{\Gamma}_0=0$, seleciona-se o modelo (36) adequado. Uma medida de plausibilidade de dois modelos M_i e M_j é fornecida pela razão das probabilidades a posteriori $\pi_{M_i}\left(.|\mathbf{Y}\right)/\pi_{M_j}\left(.|\mathbf{Y}\right)$ e o estimador MAP (maximum posterior) seleciona o modelo com maior probabilidade a posteriori. Se a probabilidade a priori satisfaz $\pi\left(M_0\right)=\pi\left(M_1\right)=...=\pi\left(M_p\right)$ deve ser selecionado o modelo com maior valor para o fator de Bayes (BF).

$$BF_{M_i}=\frac{\pi\left(M_i|\mathbf{Y}\right)}{1-\pi\left(M_i|\mathbf{Y}\right)}/\frac{\pi\left(M_i\right)}{1-\pi\left(M_i\right)} \tag{37}$$

Em geral a razão expressa pelo fator de Bayes (RBF) BF_{M_i}/BF_{M_j} é igual a razão das densidades das priori preditivas $m_i\left(\mathbf{Y}\right)/m_j\left(\mathbf{Y}\right)$ (CAO, 2010; LEE, 2007).

A abordagem de seleção de modelos requer que sejam especificadas prioris para $(\mu,\mathbf{\Gamma}_q,\mathbf{\Psi})$. Depois de determinado o submodelo M_q , é necessário fazer inferências sobre $\boldsymbol{\mu},\mathbf{\Gamma}_q$ e $\mathbf{\Psi}$ e precisamos obter as posterioris destes parâmetros. Supondo os pressupostos padrão de um modelo de AF

$$\mathbf{y}=\mathbf{\Gamma}_{\mathrm{k}}\mathbf{f}+\boldsymbol{\mu}+\mathbf{e}$$

em que $\mathbf{f}\sim N_q\left(0,\mathbf{I}\right)$ independente de $\mathbf{e}\sim N_q\left(0,\mathbf{\Psi}\right)$.

Cao (2010) postulou um teorema para o algoritmo para o amostrador de Gibbs (AGS) para AFB na abordagem de seleção de modelos. Desse teorema vale destacar nos seus pressupostos e nas prioris assumidas (“especifica como devem ser escolhidas as prioris”) sobre os parâmetros. O teorema foi apresentado incialmente por Lee (2007) e modificado por Cao

(2010). Cao (2010) salienta que a sua abordagem difere da versão de Lee (2007) porque atribui uma priori plana para $\boldsymbol{\mu}$ enquanto Lee fixa $\boldsymbol{\mu}$ para MLE (estimador de máxima verossimilhança) ($\bar{\mathbf{y}}$). No modelo de Lee $\mathbf{f} \sim N_q(0, \boldsymbol{\Phi})$ o $\boldsymbol{\Phi}$ pode assumir uma priori Whishart invertida. Enquanto que a abordagem de Cao (2010) (a mesma seguida nesta revisão) assume que os fatores comuns são independentes e identicamente distribuídos (iid). Tem se então os seguintes resultados (apresentados no Teorema 1), onde $\boldsymbol{\Gamma}_{k:q}$ representa o vetor dado para k-ésima linha de $\boldsymbol{\Gamma}_q$ e ψ_{kk} representa o k-ésimo elemento da diagonal de $\boldsymbol{\Psi}$.

Nesse teorema (que considerou-se como o primeiro) as derivações do AGS são assumidas como prioris, os pressupostos padrão do modelo de **AF**. Neste t**eorema 1** o autor postula que se assume a distribuição a priori para $\Theta = \left(\mu, \boldsymbol{\Gamma}_q, \boldsymbol{\Psi}\right)$ dado $\mu \sim 1$ que são independentes de $\left(\boldsymbol{\Gamma}_q, \boldsymbol{\Psi}\right)$, em que $\left(\Gamma_{k:q}, \psi_{kk}\right)$ são independentes para $k = 1, ..., p$ com:

$$\begin{aligned} \Gamma_{k,q} \,|\, \psi_{kk}^{-1} &\sim N_q(\Gamma_{0k}, \psi_{kk}^{-1}\mathbf{H}_{0k}) \\ \psi_{kk}^{-1} &\sim Gamma(\alpha_{0kk}, \beta_{0kk}) \end{aligned} \tag{38}$$

em que: $\mu_0, \alpha_{0k}, \beta_{0k}, \Gamma_{0k}, \mathbf{H}_{0k}$ são os hiperparâmetros. Admitindo essas condições as posterioris para os parâmetros são dadas por:

(1) $\mathbf{f}_1, ...\mathbf{f}_n$ condicionalmente independentes dado $\mathbf{Y}, \mu, \boldsymbol{\Gamma}_q, \boldsymbol{\Psi}$ com $\mathbf{f}_i \,|\left(\mathbf{Y}, \mu, \boldsymbol{\Psi}^{-1}, \boldsymbol{\Gamma}_q\right) \sim N_q\left(\mathbf{b}_i, \mathbf{B}\right)$, $\mathbf{b}_i = \mathbf{B}\boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\left(y_i - \mu\right), \mathbf{B} = \left(\mathbf{I} + \boldsymbol{\Gamma}^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{\Gamma}\right)^{-1}$;

(2) $\mu \,|\left(\mathbf{Y}, \mathbf{F}, \boldsymbol{\Psi}^{-1}, \boldsymbol{\Gamma}_q\right) \sim N_p\left(\mu_y, \boldsymbol{\Psi}/n\right)$. $\mu_y = \bar{y} - \Gamma_q \bar{\mathbf{f}}$;

(3) $\left(\psi_{11}, \Gamma_{1:q}\right), ..., \left(\psi_{pp}, \Gamma_{p:q}\right)$ condicionalmente independentes dado $\left(\mathbf{Y}, \mu, \mathbf{F}\right)$ e

$$\Gamma_{k,q} \,|\left(\mathbf{Y}, \mu, \mathbf{F}, \psi_{kk}^{-1}\right) \sim N_q\left(\mathbf{a}_k, \psi_{kk}^{-1}\mathbf{A}\right),$$

$$\mathbf{a}_k = \mathbf{A}\left[\mathbf{H}_{0k}^{-1}\Gamma_{0k} + \mathbf{F}\left(\mathbf{Y}_k - \mu_k\mathbf{1}\right)\right], \mathbf{A} = \left(\mathbf{H}_{0k}^{-1} + \mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^{-1}$$

$$\psi_{kk}^{-1} \,|\left(\mathbf{Y}, \mu, \mathbf{F}, \boldsymbol{\Gamma}_q\right) \sim Gamma\left(n/2 + \alpha_{0kk}, \beta_{kk}\right)$$

sendo que $\beta_{kk} = \beta_{0kk} + \frac{1}{2}\left[\left(\mathbf{Y}_k - \mu_k\mathbf{1}\right)^{\mathrm{T}}\left(\mathbf{Y}_k - \mu_k\mathbf{1}\right) - \mathbf{a}_k^{\mathrm{T}}\mathbf{A}^{-1}\mathbf{a}_k + \Gamma_{0k}\mathbf{H}_{0k}^{-1}\Gamma_{0k}\right]$ com $\mathbf{Y}_k^T$

Aqui não será apresentada uma demonstração para o teorema em questão. Uma prova formal pode ser encontrada em Cao (2010).

A implementação do AGS é feita de forma iterativa. O algoritmo se inicia atribuindo valores adequados de $\left(\mu, \boldsymbol{\Psi}, \boldsymbol{\Gamma}_q, \mathbf{F}\right)$, isto é, $\mu^{(0)}, \boldsymbol{\Gamma}_q^{(0)}, \boldsymbol{\Psi}^{(0)}$ e $\mathbf{F}^{(0)} = \left(\mathbf{f}_1^{(0)}, ..., \mathbf{f}_n^{(0)}\right)$, $\mu^{(1)}, \boldsymbol{\Gamma}_q^{(1)}, \boldsymbol{\Psi}^{(1)}$ e $\mathbf{F}^{(1)}$ que podem ser geradas usando as etapas (1), (2) e (3) do **teorema** 1. Cao (2010) destaca ainda que por esta abordagem requerer a especificação de prioris sobre $(\mu, \boldsymbol{\Gamma}_q, \boldsymbol{\Psi})$ é comum o uso de prioris não informativas. Isso se deve ao fato de que a escolhas comum e natural para prioris recaiam sobre priori não informativa, pois os pesquisadores nem sempre têm fortes convicções sobre o valor de um parâmetro θ. Nestas situações utilizam-se comumente o Princípio da Razão Insuficiente de Bayes - Laplace, segundo a qual todas as quantidades envolvidas no estudo têm mesma probabilidade (priori uniforme) e o método de Jeffreys baseado na medida da informação de Fisher (GELMAN et al., 2013).

A derivação do AGS das posterioris, considerando, priori não informativa para $\mu, \boldsymbol{\Gamma}_q$ **e** $\boldsymbol{\Psi}$, é apresentada por Cao (2010) (teorema 2). Segundo essa abordagem a distribuição a priori para $\boldsymbol{\Theta} = (\mu, \boldsymbol{\Gamma}_q, \boldsymbol{\Psi})$ dado $\mu \sim 1$, que são independentes de $\left(\boldsymbol{\Gamma}_q, \boldsymbol{\Psi}\right)$, em que $\left(\Gamma_{k:q}, \psi_{kk}\right)$ são independentes para $k = 1, ..., p$ com

$$\Gamma_{k:q} \,|\, \psi_{kk} \sim 1$$
$$\psi_{kk} \sim \psi_{kk}^{-1}$$

e a distribuição condicional a posteriori de $\left(\mu, \boldsymbol{\Gamma}_q, \mathbf{F}, \boldsymbol{\Psi}\right)$ é dado por

(1) $\mathbf{f}_1, ... \mathbf{f}_n$ são condicionalmente independentes dado $\mathbf{Y}, \mu, \boldsymbol{\Gamma}_q, \boldsymbol{\Psi}$ com

$$\mathbf{f}_i \,|\left(\mathbf{Y}, \mu, \boldsymbol{\Psi}^{-1}, \boldsymbol{\Gamma}_q\right) \sim N_q\left(b_i, \mathbf{B}\right), \mathbf{b}_i = \mathbf{B}\boldsymbol{\Gamma}_q^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\left(y_i - \mu\right),\ \mathbf{B} = \left(\mathbf{I} + \boldsymbol{\Gamma}_q^{\mathrm{T}}\boldsymbol{\Psi}^{-1}\boldsymbol{\Gamma}_q\right)^{-1}$$

(2)

$$\mu \,|\left(\mathbf{Y}, \mathbf{F}, \boldsymbol{\Psi}^{-1}, \boldsymbol{\Gamma}_q\right) \sim N_p\left(\mu_y, \boldsymbol{\Psi}/n\right),\ \mu_y = \bar{\mathbf{y}} - \boldsymbol{\Gamma}_q\bar{\mathbf{f}}_i$$

(3) $\left(\psi_{11}, \Gamma_{1:q}\right), ..., \left(\psi_{pp}, \Gamma_{p:q}\right)$ são condicionalmente independentes dado $\left(\mathbf{Y}, \mu, \mathbf{F}\right)$ e

$$\Gamma_{k,q} \,|\left(\mathbf{Y}, \mu, \mathbf{F}, \psi_{kk}^{-1}\right) \sim N_q\left(\mathbf{c}_k, \psi_{kk}^{-1}\left(\mathbf{F}\mathbf{F}^{\mathrm{T}}\right)^{-1}\right),\ \mathbf{c}_k = (\mathbf{F}\mathbf{F}^{\mathrm{T}})^{-1}\mathbf{F}\left(Y_k - \mu_k\mathbf{1}\right)^{\mathrm{T}}$$

$$\psi_{kk}^{-1} \,|\left(\mathbf{Y}, \mu, \mathbf{F}, \boldsymbol{\Gamma}_q\right) \sim \chi_{n-q}^2$$

em que :

$\mu_y = \bar{\mathbf{y}} - \boldsymbol{\Gamma}_q\bar{\mathbf{f}}_i$, , $\mathbf{c}_k = (\mathbf{F}\mathbf{F}^{\mathrm{T}})^{-1}\mathbf{F}\left(Y_k - \mu_k\mathbf{1}\right)^{\mathrm{T}}$ com $\mathbf{Y}_k^{\mathrm{T}}$ sendo k-ésima linha de $\mathbf{Y}$, μ_k o k-ésimo elemento de μ, 1 sendo o vetor $nx1$ de $1's$.

O procedimento de amostragem é feito pelo (AGS) de forma similar que na abordagem de concentração a posteriori**.** A abordagem de seleção de modelos requer que seja estabelecido um critério para a escolha do modelo ideal (ou seja, entre os diversos modelos em competição), pois não se conhece o melhor modelo ou que cargas deve-se reter no modelo, e o critério mais usado nesta abordagem é o fator de bayes (na seção 2.9.3 será discutido os diferentes critérios que são usados para seleção de modelos e vantagem de cada um).

2.9.3. Fator de Bayes (FB) em AFB

Para concretizar plenamente a abordagem de seleção do modelo, precisamos das probabilidades a posterioris para M_q (ou pelo menos suas razões) ou os fatores de Bayes em favor de M_q (ou pelo menos suas razões) para $q = 0,..., p$.

Suponha que se introduz uma variável latente $\mathbb{U}$ que seleciona o modelo M_i com probabilidade a priori θ_i para $i = 0,..., p$, por exemplo, $\theta_i = 1/(p+1)$. Se, adicionar esta variável latente ao Teorema 1 e 2 não muda os algoritmos de amostragem para as variáveis restantes , uma vez que $u = q$. Dado que as mudanças da dimensão do espaço do parâmetro é como mudar q, portanto, um salto reversível (reversible jump) no algoritmo Monte Carlo parece em ordem. Lee (2007) defende o uso da amostragem por caminho (path sampling) para calcular a proporção de normalização de constantes para obter os fatores de Bayes. Isto envolve computação intensiva. Segundo Cao (2010) e Lee (2007) o cálculo do fator de bayes pode ser feito usando as seguintes etapas:

Considerar dois modelos, por exemplo, M_0 e M_1, e a técnica de amostragem por importância (recomendado se os dois modelos têm a forma fechada). Então basicamente, calcula-se:

$$\begin{aligned} m(k) &= p(\mathbf{Y} \mid M_k) = \int p(\mathbf{Y}, \mathbf{F}, \boldsymbol{\Theta} \mid M_k) d\mathbf{F} d\boldsymbol{\Theta},\ k = 0,1 \\ &= \frac{\int p(\mathbf{Y}, \mathbf{F}, \boldsymbol{\Theta} \mid M_k)}{\int p(\mathbf{F}, \boldsymbol{\Theta} \mid \mathbf{Y}, M_k)} p(\mathbf{F}, \boldsymbol{\Theta} \mid \mathbf{Y}, M_k) d\mathbf{F} d\boldsymbol{\Theta} \qquad (39) \\ &= E\left(\frac{p(\mathbf{Y}, \mathbf{F}, \boldsymbol{\Theta} \mid M_k)}{p(\mathbf{F}, \boldsymbol{\Theta} \mid \mathbf{Y}, M_k)} \right), \end{aligned}$$

em que a expectância é tomado em relação $p(\mathbf{F}, \boldsymbol{\Theta} \mid \mathbf{Y}, M_k)$. Então, (39) pode ser aproximado por

$$m(k) = \frac{1}{I}\sum_{i=1}^{I} \frac{p\left(\mathbf{Y}, \mathbf{F}^{(i)}, \boldsymbol{\Theta}^{(i)} \mid M_k\right)}{p\left(\mathbf{F}^{(i)}, \boldsymbol{\Theta}^{(i)} \mid \mathbf{Y}, M_k\right)}$$

em que $\mathbf{F}^{(i)}, \boldsymbol{\Theta}^{(i)}$ é amostrada a partir da distribuição alvo $p\left(\mathbf{F}^{(i)}, \boldsymbol{\Theta}^{(i)} \mid \mathbf{Y}, M_k\right)$...por meio ... de métodos de simulação. A relação do fator de Bayes M_1 e M_0, nomeadamente, B_{10} é

$$B_{10} = \frac{m(1)}{m(0)}$$

determinado por amostragem por ponte (Bridge Sampling), sendo preferíveis para situações que M_1 e M_0, não tem forma fechada. Em vez de avaliar $m(0)$ e $m(1)$, diretamente, um modelo intermediário $M_{\frac{1}{2}}$ é construído de modo que ambos M_1 e M_0 ,sejam fechados para $M_{\frac{1}{2}}$ que um para o outro. Então, B_{10} é obtida tomando os produtos de duas razões. Além disso, melhora a eficiência e a precisão do procedimento de amostragem.

A amostragem por caminho (Path Sampling) (GELMAN; MENG, 1998), é uma extensão da amostragem por ponte, usando várias pontes na amostragem por caminho em vez de uma. Se M_1 e M_0 ,são muito distantes, uma série intermediária de modelos entre M_1 e M_0 , é construído, ditos modelos $L-1$. O fator de Bayes é equiparado ao produto de uma série de razões. Tomando o logaritmo do produto e fazendo $L \to \infty$, B_{10} é obtido. Num estudo comparativo de uma variedade de métodos para calcular o fator de Bayes, DiCiccio et al. (1997) concluíram que a amostragem por ponte proporciona tipicamente uma ordem de grandeza de melhoria sobre outros métodos no cálculo do fator de Bayes. Como a amostragem por caminho é uma generalização da amostragem por ponte, ela tem o potencial para ter ainda mais melhorias (LEE, 2007).

2.9.4. Abordagem “concentração posteriori” para análise fatorial

Métodos de verossimilhança são frequentemente utilizados para a análise de fatorial, mas estes sofrem de vários desafios computacionais. Na verdade, com restrições, por vezes, bastante arbitrárias, a serem colocados nas estimativas, pode acontecer que o software não produza respostas sensatas (CAO, 2010; LEE, 2007). Enquanto várias soluções para estes problemas têm sido propostas, como a abordagem bayesiana, discutida no tópico anterior, existe

outra abordagem muito simples que evita essas dificuldades. Esta abordagem requer uma única priori em $(\mu, \boldsymbol{\Sigma})$ e avaliar um submodelo por comparação da concentração da priori em torno do subconjunto do espaço dos parâmetros especificados pelo submodelo, com a concentração da posteriori sobre este subconjunto. Intuitivamente, se a posteriori concentra muito mais sobre este subconjunto do que a priori, então temos evidências ...por meio ... dos dados, da plausibilidade do submodelo.

Esta abordagem é conhecida como "concentração a posteriori para análise fatorial" (AF). Como referenciado acima requer que sejam especificadas prioris sobre $(\mu, \boldsymbol{\Sigma})$ apenas. Isso requer muito menos especificação da priori que a abordagem de seleção de modelos e consequentemente muito menos elicitações.

É interessante notar que a abordagem de seleção de modelos também induz uma priori sobre $(\mu, \boldsymbol{\Sigma})$, tendo uma mistura de prioris induzidas em $(\mu, \boldsymbol{\Sigma})$ pela priori de cada M_i, onde as probabilidades de mistura são as probabilidades a priori sobre os modelos (CAO, 2010; CAO; EVANS; GUTTMAN, 2010; PRESS, 1985).

Press e Shigemasu (1989) afirmam parecer impossível determinar as propriedades analíticas desta priori e relacioná-los a qualquer coisa que realmente sabemos a priori sobre $(\mu, \boldsymbol{\Sigma})$ .Considerando duas prioris sobre $(\mu, \boldsymbol{\Sigma})$, a primeira é a conjugada e é comumente usada, talvez mais conveniente do que qualquer outra. A segunda, não parece ser utilizada, devido a problemas com os cálculos (computação) das posterioris, mas parece muito mais útil. Nesta abordagem as priori assumidas são a Whishat e normal e como forma de generalização das prioris desta abordagem, Lee (2007) propõe que seja observado $\mathbf{Y} \sim (y_1, ..., y_2) \sim N(\mu, \boldsymbol{\Sigma})$ e a priori sobre $(\mu, \boldsymbol{\Sigma})$ sendo dadas por

$$\begin{aligned} \mu \mid \boldsymbol{\Sigma} &\sim N_p\left(\mu_0, \sigma_0^2 \boldsymbol{\Sigma}\right) \\ \boldsymbol{\Sigma}^{-1} &\sim Wishart\left(k_0, \mathbf{A}\right), \end{aligned} \tag{40}$$

em que:

μ_0, σ_0^2, k_0 são hiperparâmetros. As posteriori serão dadas por:

$$\begin{aligned} \mu \mid (\mathbf{Y}, \boldsymbol{\Sigma}) &\sim N\left(\mu_y, \boldsymbol{\Sigma}_y\right) \\ \boldsymbol{\Sigma}^{-1} \mid \mathbf{Y} &\sim Wishart\left(n + k_0, \mathbf{B}\right), \end{aligned} \tag{41}$$

Cao, Evans e Guttman (2010) afirmam que esta abordagem para a análise fatorial bayesiana tem várias características atraentes. Em primeiro lugar, só precisa colocar uma priori sobre o parâmetro de modelo completo (μ, Σ) em vez de uma priori em cada submodelo. Isso

reduz a necessidade de elicitação extensa e difícil, ou a imposição de prioris impróprias padrão. Além disso, é possível o uso de uma priori uniforme para matriz de correlação. Por conseguinte, só são necessários para obter prioris para os parâmetros de localização e escala que podem ser facilmente feitos de várias formas. Além disso, os autores afirmam que esta abordagem pode ser aplicada a uma série de outros problemas estatísticos.

2.9.5. Comparação de modelo (seleção de número de fatores)

Dada a importância desta decisão, foram propostos diferentes métodos para determinar o número de fatores a reter. Além disso, vários estudos têm sido realizados para avaliar a eficiência individual ou comparativa desses métodos (BUJA; EYUBOGLU, 1992). Estes estudos são em geral, uma avaliação da capacidade desses métodos para determinar o número de fatores não triviais em dados gerados por simulação. Dentre os métodos propostos merece especial destaque os estudos feitos por Akaike (1987) que propõe o critério AIC para selecionar o número apropriado de fatores para usar em um modelo de análise fatorial de MLE (máxima verossimilhança). Ele derivou a estatística para o problema de seleção de variáveis em modelos de regressão linear aninhados. A motivação era o seu desejo de controlar o efeito de supermermatrização do modelo e, portanto, de avançar para a parcimônia na modelagem estatística. Sua preocupação era com a abordagem de estimação máxima verossimilhança, pois acreditava que os modelos devem ser selecionados, ajustando o estimador de máxima verossimilhança para o número de parâmetros a ser estimado, que é exatamente o que o critério de AIC faz (PRESS; SHIGEMASU, 1999).

Mas, Akaike (1987), salienta que "a quantidade que minimiza o critério de AIC não é uma estimativa consistente do modelo correto". O procedimento é derivado como uma aproximação de amostra finita a um resultado assintótico. Schwartz (1978) também propôs um critério Bayesiano para determinar a dimensão adequada de um modelo (hierárquico), que, como foi salientado pelo Geisser e Eddy (1979), depende do tamanho da amostra, mas não sobre o tipo de erro cometido, por isso não é muito útil para a previsão (previsão dos escores fatoriais, no nosso caso). Berger e Pericchi (1996) salientam que o "BIC" parece exagerar a evidência em favor do modelo complexo, que é um" viés", se alguma coisa, vai na direção errada". Ambos AIC e BIC dependem de expansões assintóticas cujos pares dos primeiros termos dependem da função de verossimilhança (PRESS; SHIGEMASU, 1999). Zellner e Chung-Ki (1993), sugerem cinco métodos diferentes de comparação de modelos, incluindo: AIC, BIC, Quadrado Médio do Erro, o Critério de Schwarz, e o Cp Critério Mallows de 1973.

Uma suposição fundamental feita na maioria dos modelos (Bayesiano e outros) é que o número de fatores no modelo é conhecido ou pré-atribuído. Enquanto tal suposição pode ser satisfeita em alguns problemas de análise fatorial confirmatória, geralmente não é uma hipótese sustentável na análise fatorial exploratória.

Na análise Bayesiana, o fator de Bayes (RAFTERY; LEWIS, 1995) têm sido tradicionalmente usados para comparar dois modelos. No entanto, fatores de Bayes são difíceis de calcular para modelos complexos como o nosso. Nós, portanto, poderemos usar o fator pseudo-Bayes (PSBF) (GEISSER; EDDY, 1979; GELFAND, 1996) como um substituto para o fator de Bayes. Embora o fator de Bayes utilize a densidade preditiva a priori para calcular a probabilidade marginal, o PSBF baseia-se na densidade preditiva de validação cruzada dos dados. Por conseguinte, pode ser usado mesmo com prioris impróprias, que não é o nosso caso. Além disso, pode ser muito convenientemente computado para modelos de equações estruturais utilizando amostras MCMC. Ao utilizar um modelo fator analítico para inferências sobre a estrutura de covariância, ele é atraente para explicar formalmente a incerteza na escolha do número de fatores. Tem havido algum foco na literatura frequentista e bayesiana sobre o problema da seleção do número de fatores. Press e Shigemasu (1999) propõem escolher o número de fatores que têm a maior probabilidade a posteriori, notando que tal abordagem melhora os critérios comumente usados como AIC (AKAIKE, 1987) e BIC (SCHWARTZ, 1978). Para modelos hierárquicos, como modelos de fatores latentes, a justificativa do uso de BIC como uma aproximação ao fator de Bayes quebra (*breaks down*) (BERGER; GHOSH; MUKHOPADHYAY, 2003), e pode-se precisar de uma penalização diferente para modelo de maior complexidade (ZHANG; KOCKA, 2004).

Ghosh e Dunson (2008) relatam que a estimativa de probabilidades posterioris de modelos com diferentes números de fatores latentes coloca grandes desafios, tais como: (1) como escolher priores para as cargas fatoriais na lista de modelos que correspondem a números diferentes de fatores; e (2) estimar com precisão eficientemente probabilidades posterioris do modelo. Polasek (1997) considera uma abordagem para estimar probabilidades posterioris baseados em análises MCMC separada de modelos, diferindo apenas no número de fatores. Embora uma estimativa da probabilidade marginal não seja automaticamente disponível a partir da saída MCMC, um número de algoritmos tem sido proposto (CAO, 2010; MENG; WONG, 1996; ROWE, 2003).

Lopes e West (2004) propuseram um algoritmo MCMC com salto reversível (RJMCMC) (GREEN, 1995) para se mover entre os modelos com diferentes números de fatores, e realizaram uma comparação completa com estimadores de verossimilhança marginal

aproximada de análise MCMC específica em cada modelo. Lee e Song (2002) usam a abordagem de amostragem por caminho (*path sampling)* de Gelman e Meng (1998) para estimar log fatores de Bayes. Eles construíram um caminho usando um escalar $t \in [0, 1]$ para conectar dois modelos M_0 e M_1. Embora a abordagem de amostragem por caminho seja atraente por si só, a priori utilizada por Lee e Song (2002) é extremamente informativa. É um método simples de implementar e tende a ter bom desempenho em termos de acurácia. O maior problema dessa técnica para estimar o fator de Bayes, para comparar dois modelos concorrentes, é requerer a execução de algoritmos MCMC separados ao longo de uma rede correspondente a diferentes valores para uma amostragem por caminho constante. Embora essa abordagem possa potencialmente ser implementado em paralelo, no entanto, a eficiência computacional é uma preocupação, particularmente quando se tem muitos modelos diferentes sob consideração. Além disso, se as cadeias Markov individuais exibem fraca mistura o que é um problema comum em modelos com fator latente, e seria necessário executar para cada cadeia um número muito grande de iterações para se obter resultados precisos. Outro problema que é bem conhecido é que os fatores de Bayes são sensíveis à escolha de priori. Em modelos de fatores latentes, eles tendem a ser difíceis de elícitar os parâmetros, uma vez que em aplicações típicas existe uma incerteza substancial sobre os verdadeiros valores das cargas fatoriais e variâncias do erro a priori. Por isso, prefere-se escolher uma priori vaga (GOSH; DUNSON, 2008).

2.9.6. Comparação da análise fatorial bayesiana e não-bayesiana (clássica)

As vantagens de AFB são importantes. Em primeiro lugar, a informação a priori pode ser incorporada no AFB, algo que a análise fatorial frequentista não pode. Sabe-se que uma sequência de análises deve ser feita na AF afim de ter resultados satisfatórios (“seguros”). Essa informação, resultante dos estágios anteriores de análise, pode ser usada na AFB como informação a priori. Essa informação pode referir-se ao número de fatores ou as relações entre os fatores e variáveis manifestas. Outra vantagem do AFB é que o erro correlacionado pode ser facilmente incorporado no modelo sem considerável dificuldade teórica ou computacional.

Karatza (2006) listaram as vantagens da amostragem de Gibbs contra a aproximação ML (máxima verossimilhança). Eles se referem à falta de necessidade da hipótese de normalidade assintótica, da utilidade das distribuições posterioris. Este último nos fornece a chance de detectar multimodalidade e nos ajudam a inspecionar o ajuste do modelo utilizando valores-p preditivos a posteriori. Outra vantagem é o fato de que um modelo pouco identificado pode dar resultados usando priores informativas.

Kaufman e Press (1973) também enfatizam a superioridade da aplicação Bayesiana na análise fatorial. Eles apoiam a ideia de que as restrições em análise fatorial clássica são muito "dogmáticas" e as matrizes de carga resultantes não são únicas, uma vez que podem ser alteradas por uma rotação adequada. Isso não acontece no AFB, que só precisa de especificação cuidadosa de distribuições a priori. Outro problema da análise frequentista em relação à bayesiana é que os resultados da análise clássica devem ser interpretados com cautela, uma vez que casos Heywood (PRESS, 1989; SMITH; CULLIS; THOMPSON, 2001) podem ocorrer durante a avaliação. Uma grande vantagem da análise bayesiana é a prevenção de casos Heywood. Isso nos ajuda a obter estimativas estáveis dos parâmetros do modelo.

Na abordagem bayesiana propomos usar a parametrização da decomposição espectral diferente da triangular inferior da matriz proposta por Lopes e West (2004).

2.10. Inferência bayesiana no estudo da interação GE

A análise de conjuntos de dados MET, visando o estudo da interação GE, tem figurado como principal objetivo em programas de melhoramento genético de plantas. Apesar de vários métodos terem sido propostos para esse fim, o uso da inferência bayesiana ainda tem sido limitado, como observado por Crossa et al. (2011). Não obstante a escassa literatura envolvendo a análise de dados MET dentro da ótica bayesiana, alguns trabalhos podem ser destacados como as abordagens propostas por Cotes et al. (2006), Edwards e Jannink (2006), Orellana, Edwards e Carriquiry (2014) e Theobald, Talbot e Nabugoomu (2002).

Theobald, Talbot e Nabugoomu (2002), por exemplo, utilizaram a inferência bayesiana para estudar conjuntos de dados desbalanceados em ensaios MET com um modelo hierárquico. Edwards e Jannink (2006) propuseram o modelo hierárquico que considera a heterogeneidade de variância da interação genótipo x ambiente e residual e discutiram as propriedades do estimador Bayesiano obtido a partir do modelo considerando variâncias heterogêneas da GEI, bem como variâncias heterogêneas residuais. Verificaram que o estimador bayesiano é uma média ponderada, onde os pesos são dados por funções do tamanho da amostra e da repetibilidade. Para locais com amostra de tamanho grande e /ou repetibilidade, o estimador vai colocar mais peso na média para esse local, mostrando que este local é mais “confiável”. Se o tamanho da amostra e repetibilidade diminuir o estimador vai colocar mais peso sobre a priori, na ausência de uma priori informativa o peso vai ser colocada na média geral, os autores referem a isso como uma “penalidade” para instabilidade de cultivar.

Orellana (2012) usando inferência bayesiana para estudar a presença de heterogeneidade de variâncias tanto na GEI quanto residual encontraram evidências muito convincentes de

heterogeneidade de variância nos dois níveis. E seus resultados sugerem que o uso de um modelo que leva em conta a heterogeneidade, tanto de variâncias GEI quanto residual pode levar a diferentes estimativas de efeitos genéticos. Verificaram ainda que a quantidade de heterogeneidade das variâncias reveladas pelos resultados, mesmo nos casos em que se esperava uma variância muito pequena, demonstra que, ao utilizar o modelo homogêneo praticante se está violando uma hipótese muito forte (na análise).

No contexto dos modelos lineares-bilineares, como os populares modelo de efeitos principais aditivos e interação multiplicativa (AMMI) e o de efeitos de genótipos mais interação GEI (GGEbiplot), a complexa estrutura paramétrica dos termos que descrevem a interação, imposta pela decomposição por valores singulares da matriz de resíduos do ajuste aos efeitos principais, tem imposto grandes dificuldades para o processo de amostragem, principalmente no que se refere aos vetores singulares, cujo suporte parta a distribuição conjunta a posteriori não é trivial.

Essas dificuldades foram contornadas por Viele e Srinivasan (2000) - os primeiros a tratarem o modelo AMMI em uma abordagem bayesiana. Os referidos autores propuseram a modelagem da interação GE utilizando distribuições esféricas para os vetores singulares e normais truncadas para os valores singulares. Com a utilização de técnicas MCMC conduziram a amostragem das coordenadas dos vetores singulares via transformação linear ortogonal, em que os vetores eram primeiro amostrados no subespaço corrigido e depois transformados um a um para o subespaço correto pelo método de ortogonalização de Gram-Schmidt. Utilizando essa manobra algébrica conseguiram preservar de forma eficiente as restrições do modelo AMMI. A abordagem de modelos lineares-bilineares, no contexto de dados direcionais, tem oferecido uma maneira elegante e eficiente de analisar a interação GEI. A distribuição uniforme esférica é um caso particular da distribuição de Von Mises-Fisher (MARDIA; KENT; BIBBY, 1979) e foi utilizada com originalidade por uma priori por Viele e Srinivasan (2000) para estimar os parâmetros de interação do Modelo AMMI.

Mais recentemente Crossa et al. (2011) e Perez-Elizald, Jarquin e Crossa (2011), a luz dos trabalhos de Liu (2001) e Viele e Srinivasan (2000) utilizaram o modelo AMMI-bayesiano para superar limitações da abordagem frequentista como, por exemplo, a incorporação de inferência ao biplot, que pelo método tradicional tem recebido inúmeras críticas colocando sua validade em cheque (YAN et al., 2010; YANG et al., 2009). Perez-Elizald, Jarquin e Crossa (2011) mostraram como incorporar informações históricas de experimentos que podem ser incorporadas a um modelo com notação matricial considerando as distribuições posterioris de

matrizes ortonormais que surgem na análise de dados multivariados, como sugerido por Hoff (2009).

Os autores de Oliveira et al. (2015) e Silva et al. (2015) mostraram que além da grande flexibilidade para incorporar inferência paramétrica ao biplot, o método bayesiano permite, entre outras coisas, atribuir efeito aleatório para genótipos, bem como um método parcimonioso para a seleção do número de termos bilineares do modelo AMMI. da Silva et al. (2015), apresentou uma justificativa bayesiana para o método Shrinkage proposto por Cornelius e Crossa (1999), bem como as vantagens sugestivas de que a abordagem bayesiana oferece em detrimento aos métodos frequentistas. Josse et al. (2014), por sua vez, apresentaram uma abordagem bayesiana para o modelo AMMI, de forma a contornar as dificuldades de se amostrar os parâmetros bilineares na perspectiva dos dados direcionais. Jarquín et al. (2016) apresentaram recentemente uma proposta bayesiana para o modelo SREG, usando abordagem multinível (hierárquica) para o estudo da interação genótipo x ambiente x ano, mostraram ainda que a análise Bayesiana hierárquica é adequada para lidar com as diferenças entre períodos de avaliação, porque todas as informações estão disponíveis (informação a priori e informações de cada período de avaliação) e poderiam potencialmente ser incluídos na análise em simultâneo. Os autores defendem ainda que por este modelo incorporar naturalmente todas as informações disponíveis períodos (anos), a distribuição posteriori correspondente tem uma predição do parâmetro que melhora a sua precisão com o aumento número de períodos.

Apesar das utilidades e propostas apresentadas sobre a extensão dos modelos AMMI e SREG sob a ótica bayesiana, é crescente uso dos modelos FA na análise de dados de ensaios multi-características e MET. Isto talvez porque os métodos bayesianos propostos até hoje assumem uma estrutura de variância homogênea, seja genética ou residual, quer no melhoramento animal ou vegetal, cuja implementação? e uso não é trivial devido as grandes restrições que devem ser assumidos a priori e falta de identificabilidade para estimação dos parâmetros do modelo, e ocorrência de casos Heywood. Diante dessas limitações Campos e Gianola (2007) apresentaram uma abordagem bayesiana para o modelo FA com aplicação em dados de melhoramento animal. Nessa proposta os autores mostraram como deve ser feita amostragem dos parâmetros do modelo e apresentaram as condicionais completas a posteriori dos parâmetros. O método proposto por esses autores é muito útil, pois permite a eliminação de imposição de restrições e identificabilidade do modelo, que torna o uso desses modelos em abordagem frequentista restritiva fora dos programas estatísticos concebidos para o efeito como GenStat e Asreml (que não são de código aberto). Porém, apesar da utilidade dessa proposta, várias questões ficaram por ser resolvidas, tais como fazer a amostragem direta dos parâmetros

sem a necessidade de amostrar primeiros os parâmetros do modelo misto, como usar o modelo em situação de dados desbalanceados e como esta proposta seria útil no melhoramento vegetal para o estudo da interação GE. Nesta tese procuramos responder algumas destas questões, além de propor uma forma diferente de tratar o modelo FA, por meio da decomposição espectral.

3. MATERIAL E MÉTODOS

Nesta seção, serão apresentados o material experimental utilizado neste trabalho e os métodos que usados para as análises realizadas.

3.1 Material

3.1.1 Dados simulados

Foi simulado um conjunto de dados com 20 genótipos (G1-G20), avaliados em cinco ambientes (E1-E5), usando o delineamento de blocos ao acaso com duas repetições.

Foram simulados cinco genótipos com interação positiva em todos ambientes porem instáveis, cinco genótipos com interação negativa e instáveis e 10 genótipos com interações positivas e negativas e estáveis. No caso dos genótipos instáveis, a variância adotada foi dois e nos estáveis a variância foi um. Para dados foi estimada uma herdabilidade (h^2) média como razão entre a variância genética pela soma entre a variância genética média e variância residual.

3.1.2 Dados experimentais

Os dados experimentais foram descritos por Melo et al. (2014). Esses dados referem à produção ($kg.ha^{-1}$) de 50 híbridos de milho (G1-G50), avaliados em 10 locais (E1-E10), no inverno de 2010, no campo experimental de Uberlândia. O experimento foi conduzido no modelo experimental de blocos incompletos (DBI) com duas repetições e cada parcela possuía quatro linhas de 5 m com espaçamento entre linhas de 70 cm.

3.2 Método

3.2.1 Modelo estatístico

Análise de Fatores sob modelos mistos multiplicativos usando abordagem Bayesiana.

O modelo misto multivariado, em notação matricial, é apresentado como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\mathbf{u} + \boldsymbol{\varepsilon} \quad , \tag{42}$$

em que $\mathbf{y}(n\times 1)$ é o vetor de observações para p ambientes, q blocos e m genótipos e os vetores $\boldsymbol{\beta}(\mathrm{p}\,q\times 1)$, $\mathbf{u}(mp\times 1)$ e $\boldsymbol{\varepsilon}(n\times 1)$ são os vetores de efeitos de fixos, efeitos aleatórios e residuais, respectivamente. As expressões $\mathbf{X}(n\times pq)$, $\mathbf{Z}(n\times mp)$, por sua vez, denotam as matrizes de delineamentos referentes à $\boldsymbol{\beta}$, e $\mathbf{u}$ respectivamente. Por simplicidade, assume-se que $\mathbf{u}$ representa os efeitos genéticos aditivos.

Assumindo que o modelo (42) é de posto completo p e que a matriz de covariâncias $\mathbf{\Sigma}$ é representada por uma estrutura FA ($\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Gamma}^{\top} + \mathbf{\Psi}$), o modelo (42) é obtido por meio do ajuste dos fatores comuns e específicos separadamente $\left[\mathbf{u} = \left(\mathbf{\Gamma}_{\mathrm{p}} \otimes \mathbf{I}_{\mathrm{m}}\right)\mathbf{f} + \boldsymbol{\delta}\right]$ e pode ser reescrito como:

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{Z}\left(\mathbf{\Gamma}_{\mathrm{p}} \otimes \mathbf{I}_{\mathrm{m}}\right)\mathbf{f} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon} \ . \tag{43}$$

O modelo assim obtido é designado modelo misto fatorial analítico ou simplesmente modelo fator analítico, com $\mathbf{\Gamma}(\mathrm{p} \times k)$, $\mathbf{f}\,(\mathrm{m}k \times 1)$ e $\boldsymbol{\delta}\,(mp \times 1)$ são vetores de efeitos comuns genéticos aditivos específicos, respectivamente, $\boldsymbol{\varepsilon}$ é o vector de erros, $\mathbf{X}$ a matriz de incidência referentes a $\boldsymbol{\beta}$, e $\mathbf{Z}$ matriz de incidência referente a $\mathbf{f}$ e $\boldsymbol{\delta}$, onde k representa o número de termos multiplicativos.

Dado que $\mathbf{\Gamma}\mathbf{\Gamma}^{\top}$ é uma matriz simétrica podemos reparametrizar o modelo (43) observando que a matriz de cargas fatoriais pode ser obtida por $\mathbf{\Gamma} = \mathbf{\Lambda}^{\frac{1}{2}}\mathbf{V}$. Nessa expressão $\mathbf{V}$ representa a matriz de vetores singulares e $\mathbf{\Lambda}$ uma matriz diagonal formada pelos autovalores obtidos pela decomposição espectral da matriz de cargas. A decomposição espectral da matriz de covariância pode ser representada da seguinte forma quando se ajustam todos os termos do modelo:

$$\mathbf{\Sigma} = \mathbf{\Gamma}\mathbf{\Gamma}^{\top} = \mathbf{V}_{\mathrm{p}}\mathbf{\Lambda}_{\mathrm{p}}\mathbf{V}_{\mathrm{p}}^{\top} = \sum_{i=1}^{p} \eta_i a_i \alpha_i^{\top} \tag{44}$$

Um modelo com $p = k$ (k sendo o posto da matriz) termos multiplicativos é dito ser de posto completo, de forma que os efeitos específicos são assumidos ausentes. Contudo, a vantagem em se utilizar a análise de fatores ocorre para k bem menor do que p , de forma que o número de parâmetros na análise de fatores, $k\left(p+1\right)$, se torna bem menor do que aqueles $p\left(p+1\right)/2$ parâmetros de $\mathbf{\Sigma}$.

Utilizando componentes principais para descrever as cargas fatoriais e considerando as propriedades da decomposição espectral e transformações lineares apropriadasé possível reescrever o termo da equação (43) que envolve as cargas como se segue:

$$\mathbf{ZL} = \mathbf{Z}\left(\mathbf{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}}\right) = \sum_{k=1}^{p} \lambda_k \mathit{diag}\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1 \tag{45}$$

Mais detalhes sobre estas propriedades e demonstrações do teorema podem ser encontrados em Songgui e Suju (2002) e no apêndice-A.

Substituindo o produto de Kronecker em (43) pelo somatório da expressão (45) o modelo FA pode ser expresso por:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}; \tag{46}$$

No modelo (46) λ_k e $\boldsymbol{\alpha}_k$ são os *k*-ésimos valores singulares e vetor singular da decomposição espectral da matriz $\boldsymbol{\Sigma}$,as matrizes $\mathbf{X}_1$, $\mathbf{X}_2$, $\mathbf{Z}$ e $\mathbf{Z}_1$ são matrizes de delineamentos destinadas a distribuir os efeitos para as parcelas no modelo. Reescrevendo o modelo como (46) por meio do somatório, conseguimos distribuir os efeitos de forma eficiente e evitamos trabalhar com o produto de Kronecker que acarreta maiores dificuldades em termos computacionais e manipulações algébricas. A expressão apresentada em (46), por outro lado, oferece maior simplicidade, velocidade computacional e facilidade para manipulações algébricas, evitando o uso de artifícios, como trabalhar com o produto de Hadamart e Kronecker, por exemplo.

Para $p = k$, (46) é um modelo equivalente ao (43). Caso contrário, ou seja, para $k < p$, considerando os primeiros k componentes principais somente, este modelo terá dimensão reduzida, pois passa a ter um ruído que deveria ser explicado pelos demais componentes.

Dessa forma, o modelo apresentado em (46) é mais tratável de ponto de vista bayesiano em relação ao seu equivalente (43). A distribuição condicional dos dados tem densidade normal multivariada:

$$\mathbf{y} \mid \boldsymbol{\beta}, \lambda, \boldsymbol{\alpha}, \mathbf{f}, \mathbf{R} \sim \mathbf{N}[\mathbf{X}_1\boldsymbol{\beta} + \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}], \mathbf{I} \otimes \mathbf{R}].$$

3.2.2 Implementação da análise bayesiana do modelo proposto.

O processo de estimação Bayesiana do modelo proposto é baseado nas distribuições a posteriori condicionais para todos os parâmetros.

3.2.2.1 Distribuição a priori (prioris independentes)

A distribuição a priori é parte chave da inferência bayesiana e representa as informações sobre a incerteza do parâmetro que é combinado com a distribuição de probabilidade dos dados para se obter a distribuição a posteriori, o que por sua vez, é usado para inferências futuras e decisões que envolvem o parâmetro. Neste estudo foram considerados os pressupostos do

modelo de análise de fatores sobre os modelos mistos via máxima verossimilhança restrita (REML) como prioris para o modelo (46). Neste caso, as distribuições prioris para o modelo FA (46) atribuídos aos parâmetros foram como se segue:

i. $\boldsymbol{\beta}|...\sim {}^{1}\!\!\diagup\!\!{}_{\sigma}$;

ii. $\mathbf{f}|...\sim N\left[\mu_{\mathrm{f}_k},\mathbf{I}_{\mathrm{m}}\right],\mu_{\mathrm{f}_k}\to\mathbf{0}$, $\mathbf{I}_{\mathrm{m}}$ -é uma matriz identidade com os elementos independentes;

iii. $\boldsymbol{\delta}|...\sim N\left[\mathbf{0},\boldsymbol{\Psi}\right]$, em que: $\boldsymbol{\Psi}=diag\left(\psi_{11},...,\psi_{kk}\right),k=1,...,\mathrm{p}$;

iv. $\mathbf{R}|...\sim IW_p\left(S\right)$, em que: $\mathbf{R}=diag\left(\sigma^2_{e_1},...,\sigma^2_{e_k}\right),k=1,...,\mathrm{p}$ e para $\sigma^2_{e_k}$ cada $Inv-\chi^{-2}\left(\nu_e,S^2_k\right),\nu_e=0,S^2_k=0,$ dado $\mathbf{R}$ ser uma estrutura diagonal;

v. $\boldsymbol{\alpha}_k|\sim...\text{uniforme esférica no espaço corrigido}$;

vi. $\boldsymbol{\lambda}_k|...\sim N^{+}\left[\mu_{\lambda_k},\sigma^2_{\lambda_k}\right],\ \mu_{\lambda_k}=0;\ \sigma^2_{\lambda_k}=\infty$;

vii. $\boldsymbol{\Psi}=diag\left(\psi_{11},...,\psi_{kk}\right)\ e\ k=1,..,\mathrm{p}$, portanto; $\psi_{kk}|...\sim \text{Inv}-\chi^{-2}\left(\nu_k,S^2_k\right),\nu_k=0,\ S^2_k=0;$

em que, N indica a distribuição normal multivariada, N^{+} é a distribuição normal truncada positiva, Inv$-\chi^{-2}$ é a distribuição qui-quadrado escalada invertida. Todas as prioris apresentadas satisfazem as restrições do modelo e são priores condicionalmente conjugadas. A distribuição uniforme esférica é um caso especial de uma distribuição de Von Mises-Fisher (MARDIA; KENT; BIBBY, 1979), que é uma família de conjugadas. Se as prioris para os parâmetros são informativas ou não-informativas depende dos seus parâmetros; para priores normais, variância infinita torna as prioris quase não-informativas.

3.2.2.2 Distribuições a posteriori condicionais completas para os parâmetros

Assumindo que a verossimilhança dos dados de (46) é dada por:

$$\mathrm{L}(\mathrm{y}|\boldsymbol{\varphi})=\frac{1}{2\pi\left|R\right|^{-1}}\exp\left\{-\frac{1}{2}\left(\mathbf{y}-\boldsymbol{\theta}\right)^{\top}\mathbf{R}^{-1}\left(\mathbf{y}-\boldsymbol{\theta}\right)\right\} \quad (47)$$

em que: $\boldsymbol{\varphi}=\left(\boldsymbol{\beta},\boldsymbol{\lambda},\boldsymbol{\delta},\boldsymbol{\alpha},\mathbf{f},\mathbf{R},\boldsymbol{\Psi}\right)$; $\boldsymbol{\theta}=\mathbf{X}_1\boldsymbol{\beta}+\sum_{k=1}^{m}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k+\mathbf{Z}\boldsymbol{\delta}$;

Aplicando o teorema de Bayes após a elicitação de todos os parâmetros, considerando o modelo apresentado em (46), a posteriori conjunta é dada por:

$$p(\boldsymbol{\varphi}\,|\,\mathbf{y}) \propto \mathrm{L}(\mathrm{y}\,|\,\boldsymbol{\varphi})\,p(\boldsymbol{\beta})\,p(\lambda)\,p(\boldsymbol{\alpha})\,p(\boldsymbol{\delta})\,p(\mathbf{f})\,p(\mathbf{R})\,p(\boldsymbol{\Psi}) \quad (48)$$

As distribuições condicionais posterioris são como se segue:

i) Distribuição condicional completa a posteriori para $\boldsymbol{\beta}$

Em (46) fazendo $\mathbf{t} = \mathbf{y} - \sum_{k=1}^{m} \lambda_k \, diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)\mathbf{X}_2\mathbf{f}_k - \mathbf{Z}\boldsymbol{\delta}$

A distribuição a posteriori de $\boldsymbol{\beta}$ é dada por:

$$p(\boldsymbol{\beta}\,|\,...) \propto \exp\left\{\left(\mathbf{t}-\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}\left(\mathbf{t}-\mathbf{X}_1\boldsymbol{\beta}\right)\right\}\exp\left(\frac{1}{\sigma_\beta^2}\boldsymbol{\beta}^{\top}\boldsymbol{\beta}\right)$$

$$p(\boldsymbol{\beta}\,|\,...) \propto \exp\left\{\left(t-\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}\left(t-\mathbf{X}_1\boldsymbol{\beta}\right)+\frac{1}{\sigma_\beta^2}\boldsymbol{\beta}^{\top}\boldsymbol{\beta}\right\} \quad (49)$$

Resolvendo a parte do exponencial temos:

$$\left(t-\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}\left(t-\mathbf{X}_1\boldsymbol{\beta}\right)+\frac{1}{\sigma_\beta^2}\boldsymbol{\beta}^{\top}\boldsymbol{\beta} = t^{\top}\mathbf{R}^{-1}t-2\left(\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}t+\left(\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}\mathbf{X}_1\boldsymbol{\beta}+\frac{1}{\sigma_\beta^2}\boldsymbol{\beta}^{\top}\boldsymbol{\beta} =$$

$$= t^{\top}\mathbf{R}^{-1}t-2\left(\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}t-\boldsymbol{\beta}^{\top}\left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1+I\frac{1}{\sigma_\beta^2}\right)\boldsymbol{\beta}$$

Assumindo que a solução (estimativa de máxima verossimilhança) dos betas do modelo (43) são

$$\hat{\boldsymbol{\beta}} = \left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{t} = \left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^{\top}\mathbf{R}^{-1}\left\{\mathbf{y}-\left(\sum_{k=1}^{m}\lambda_k\,diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k+\mathbf{Z}\boldsymbol{\delta}\right)\right\} \quad (50)$$

Poderemos somar zero a expressão (49), obtendo

$$t^{\top}\mathbf{R}^{-1}t-2\left(\mathbf{X}_1\boldsymbol{\beta}\right)^{\top}\mathbf{R}^{-1}t-\boldsymbol{\beta}^{\top}\left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1+I\frac{1}{\sigma_\beta^2}\right)\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}^{\top}\left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1+I\frac{1}{\sigma_\beta^2}\right)\hat{\boldsymbol{\beta}}+\hat{\boldsymbol{\beta}}^{\top}\left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1+\mathbf{I}\frac{1}{\sigma_\beta^2}\right)\hat{\boldsymbol{\beta}}$$

Então a forma quadrática do beta pode ser combinada e a expressão acima ser escrita como:

$$\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)^{\top}\left(\mathbf{X}_1^{\top}\mathbf{R}^{-1}\mathbf{X}_1+I\frac{1}{\sigma_\beta^2}\right)^{-1}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right) \quad (51)$$

e, por meio de manipulações algébricas, assumindo que $\frac{1}{\sigma_\beta^2} \to 0$, e devolvendo (51) em (49) temos:

$$p(\boldsymbol{\beta}|...) \propto \exp\left\{-\frac{1}{2\sigma_e^2}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)^\top \left(\mathbf{X}_1^\top \mathbf{R}^{-1}\mathbf{X}_1\right)^{-1}\left(\boldsymbol{\beta}-\hat{\boldsymbol{\beta}}\right)\right\}$$

Logo a distribuição posteriori condicional para os betas é:

$$\boldsymbol{\beta}|... \sim N\left(\left(\mathbf{X}_1^\top \mathbf{R}^{-1}\mathbf{X}_1\right)^{-1}\mathbf{X}_1^\top \mathbf{R}^{-1}\left\{\mathbf{y}-\left(\sum_{k=1}^{p}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}\right)\right\}, \left(\mathbf{X}_1^\top \mathbf{R}^{-1}\mathbf{X}_1\right)^{-1}\right) \quad (52)$$

ii) Distribuição condicional completa a posteriori para λ_k

Considerando que os escores fatoriais e os valores singulares podem ser representados por:

$\mathbf{f} = (\mathbf{f}_1^\top,...,\mathbf{f}_j^\top)^\top$; $\lambda = (\boldsymbol{\lambda}_1^\top,...,\boldsymbol{\lambda}_j^\top)^\top$

Poderemos escrever (43) como:

$$\sum_{k=1}^{p}\lambda_k diag\left(\mathbf{Z}\boldsymbol{\alpha}_k\right)\mathbf{X}_2\mathbf{f}_k = \lambda_1 diag\left(\mathbf{X}_2\boldsymbol{\alpha}_1\right)\mathbf{Z}_1\mathbf{f}_1 + ... + \lambda_p diag\left(\mathbf{X}_2\boldsymbol{\alpha}_p\right)\mathbf{Z}_1\mathbf{f}_p \quad (53)$$

Sejam $A_{1k} = \mathbf{y} - \left[\mathbf{X}_1\beta + \sum_{k'\neq k}^{p}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}\right]$ e $A_{2k} = diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k$

$$p(\lambda_k|...) \propto \exp\left\{-\frac{1}{2}\left(\mathbf{A}_{1k}-\lambda_k\mathbf{A}_{2k}\right)^\top \mathbf{R}^{-1}\left(\mathbf{A}_1-\lambda_k\mathbf{A}_{2k}\right)\right\}$$

$$p(\lambda_k|...) \propto \exp\left\{-\frac{1}{2}\left[\left(\mathbf{A}_{1k}-\lambda_k\mathbf{A}_{2k}\right)^\top \mathbf{R}^{-1}\left(\mathbf{A}_{1k}-\lambda_k\mathbf{A}_{2k}\right)\right]\right\}$$

$$p(\lambda_k|...) \propto \exp\left\{-\frac{1}{2}\left[A_{1k}^\top\mathbf{R}^{-1}A_{1k} - A_{1k}^\top\mathbf{R}^{-1}\lambda_k A_{2k} - \lambda_k A_{2k}^\top\mathbf{R}^{-1}A_{1k} + \lambda_k^2 A_{2k}^\top\mathbf{R}^{-1}A_{2k}\right]\right\}$$

$$p(\lambda_k|...) \propto \exp\left\{-\frac{1}{2}\left[\lambda_k^2\left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{2k}\right) - \mathbf{A}_{1k}^\top\mathbf{R}^{-1}\lambda_k\mathbf{A}_{2k} - \lambda_k\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{1k}\right]\right\}$$

$$p(\lambda_k|...) \propto \exp\left\{-\frac{1}{2}\left[\left(\lambda_k - \left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{2k}+\right)^{-1}\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{1k}\right)^\top\left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{2k}\right)\left(\lambda_k - \left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}A_{2k}\right)^{-1}\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{1k}\right)\right]\right\}$$

$$\lambda_k|... \propto N^+\left[\left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{2k}\right)^{-1}\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{1k}, \left(\mathbf{A}_{2k}^\top\mathbf{R}^{-1}\mathbf{A}_{2k}\right)^{-1}\right] \quad (54)$$

iii) Distribuição condicional completa para $\boldsymbol{\alpha}_k$

Assumindo:

$$\boldsymbol{\theta} = \mathbf{X}_1\boldsymbol{\beta} + \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}$$

$$\mathbf{A}_{4k} = \mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} + \sum_{k'\neq k}^{p} \lambda_{k'} diag\left(\mathbf{X}_2\alpha_{k'}\right)\mathbf{Z}_1\mathbf{f}_{k'} + \mathbf{Z}\boldsymbol{\delta}\ ;\ \Delta_k = \lambda_k diag\left(\mathbf{Z}_1\mathbf{f}_k\right)\mathbf{X}_2$$

A distribuição condicional para α_k pode ser obtida como segue abaixo:

$$p\left(\alpha_k \mid ...\right) \propto \exp\left\{-\frac{1}{2}\left(y-\boldsymbol{\theta}\right)^{\top}\mathbf{R}^{-1}\left(y-\boldsymbol{\theta}\right)\right\} \tag{55}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\mathbf{A}_{4k}-\Delta_k\alpha_k\right)^{\top}\mathbf{R}^{-1}\left(\mathbf{A}_{4k}-\Delta_k\alpha_k\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\mathbf{A}_{4k}^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k} - \mathbf{A}_{4k}^{\top}\mathbf{R}^{-1}\Delta_k\alpha_k - \alpha_k^{\top}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k} + \alpha_k^{\top}\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\alpha_k\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(2\alpha_k^{\top}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k} + \alpha_k^{\top}\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\alpha_k\right)\right\}$$

$$\propto \exp\left\{-\frac{1}{2}\left(\left[\alpha_k - \left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\left[\alpha_k - \left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]\right)\right\}$$

$$\boldsymbol{\alpha}_k \mid ... \sim N\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}, \left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\right] \tag{56}$$

Embora as prioris para $\boldsymbol{\alpha}_k$ sejam uniformes esféricas não será possível obter uma Von Mises-Fisher como condicional conjunta a posteriori devido à heterocedasticidade assumida no modelo. Logo a condicional, como veremos, será uma normal multivariada o que obviamente viola as restrições do modelo FA. É importante observar que se $\mathbf{R} = \mathbf{I}\sigma^2$ a distribuição para $\boldsymbol{\alpha}_k$ seria Von Misses-fisher.

Vimos então que $\boldsymbol{\alpha}_k$ segue uma distribuição normal multivariada m dimensional. Contudo, os vetores singulares, assim gerados, violam a restrição de norma unitária (e devem ser normalizados). Além disso, o suporte da posteriori conjunta para garantir a restrição de ortogonalidade não é trivial e a amostragem em $\Re^m$ seria um processo complicado.

Para contornar essas dificuldades, a amostragem será realizada no espaço corrigido, livres da restrição de ortogonalidade, por variáveis auxiliares (que serão definidas mais adiante)

e por meio de transformações lineares ortogonais serão devolvidos no correto subespaço em $\Re^m$ com as restrições impostas pelo modelo FA.

iv) Distribuição condicional para os escores fatoriais $\mathbf{f}$

Assumindo que: $\mathbf{f} = (\mathbf{f}_1^\top, ..., \mathbf{f}_m^\top)^\top$ e $\boldsymbol{\lambda} = (\boldsymbol{\lambda}_1^\top, ..., \boldsymbol{\lambda}_m^\top)^\top$ temos que a priori para $\mathbf{f}$ é:

$$p(\mathbf{f}) \propto \exp\left(-\frac{1}{2}\mathbf{f}^\top\mathbf{f}\right).$$

Assumindo que a amostragem será feita para o $\mathbf{f}$ poderemos rescrever o (45) como:

$$\mathbf{A}_3 = \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1 = \mathbf{ZL} = \mathbf{W}$$

Combinando a verossimilhança e a priori, a distribuição condicional a posteriori de $\mathbf{f}$ é dada por:

$$p(\mathbf{f}|...) \propto |R|^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \mathbf{A}_3\mathbf{f} - \mathbf{Z}\boldsymbol{\delta}\right)^\top \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \mathbf{A}_3\mathbf{f} - \mathbf{Z}\boldsymbol{\delta}\right) + \mathbf{f}^\top\mathbf{f}\right\}$$

com:

$$\varphi = \mathbf{y} - \mathbf{X}_1\beta - \mathbf{Z}\boldsymbol{\delta}.$$

Assim temos:

$$p(\mathbf{f}|...) \propto \exp\left\{-\frac{1}{2}\left(\varphi - \mathbf{A_3}\mathbf{f}\right)^\top \mathbf{R}^{-1}\left(\varphi - \mathbf{A_3}\mathbf{f}\right)\right\}\exp\left(\mathbf{f}^\top\mathbf{f}\right)$$

$$p(\mathbf{f}|...) \propto \mathbf{R}^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\left(\varphi^\top\mathbf{R}^{-1}\varphi - 2\mathbf{A_3}^\top\mathbf{R}^{-1}\varphi + \mathbf{f}^\top\mathbf{A_3}^\top\mathbf{R}^{-1}\mathbf{A_3} + \mathbf{f}^\top\mathbf{f}\right)\right\}$$

Por meio de manipulações algébricas chegamos que condicional dos escores fatoriais é a uma normal multivariada.

$$\mathbf{f}|... \sim \mathbf{N}\left(\left(\mathbf{I} + \mathbf{A}_3^\top\mathbf{R}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3^\top\mathbf{R}^{-1}\varphi;;\left(\mathbf{I} + \mathbf{A}_3^\top\mathbf{R}^{-1}\mathbf{A}_3\right)^{-1}\right) \quad (57)$$

v) Distribuição condicional para variâncias específicas $\boldsymbol{\delta}$

Considerando a priori atribuída a variância específica e o modelo (46) fazendo $\mathbf{A}_6 = \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k$ e

$\partial = \mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \mathbf{A}_6$;

$\mathbf{V} = \mathbf{I}_m \otimes \mathbf{R}$;

e $\mathbf{M} = \mathbf{I}_m \otimes \boldsymbol{\Psi}$; Aplicando o teorema de Bayes temos:

$$p(\boldsymbol{\delta}|...) \propto \exp\left\{-\frac{1}{2}(\partial - \mathbf{Z}\boldsymbol{\delta})^\top \mathbf{V}^{-1}(\partial - \mathbf{Z}\boldsymbol{\delta}) + \boldsymbol{\delta}^\top \mathbf{M}^{-1}\boldsymbol{\delta}\right\} \quad (57)$$

Resolvendo a parte exponencial de (57) temos:

$$\partial^\top \mathbf{V}^{-1}\partial - \partial^\top \mathbf{V}^{-1}\mathbf{Z}\boldsymbol{\delta} - \boldsymbol{\delta}^\top \mathbf{Z}^\top \mathbf{V}^{-1}\partial + \boldsymbol{\delta}^\top \mathbf{Z}^\top \mathbf{V}^{-1}\mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\delta}^\top \mathbf{M}^{-1}\boldsymbol{\delta} = \quad (58)$$

$$= \boldsymbol{\delta}^\top \left(\mathbf{M}^{-1} + \mathbf{Z}^\top \mathbf{V}^{-1}\mathbf{Z}\right)\boldsymbol{\delta} - \partial^\top \mathbf{V}^{-1}\mathbf{Z}\boldsymbol{\delta} + \left[\partial^\top \mathbf{V}^{-1}\mathbf{Z}\boldsymbol{\delta}\right]^\top + \partial^\top \mathbf{V}^{-1}\partial$$

$$Se\ B_1 = \left(\mathbf{M}^{-1} + \mathbf{Z}^\top \mathbf{V}^{-1}\mathbf{Z}\right)^{-1};\ b_1 = \mathbf{B}_1 \mathbf{Z}^\top \mathbf{V}^{-1}\partial^\top \text{, então } \mathbf{b}_1^\top \mathbf{B}_1^{-1} = \partial^\top \mathbf{V}^{-1}\mathbf{Z}$$

Assim que temos (58) fica:

$$(\boldsymbol{\delta} - \mathbf{b}_1)^\top B_1^{-1}(\boldsymbol{\delta} - \mathbf{b}_1) + c \quad (59)$$

Então (57) fica:

$$p(\boldsymbol{\delta}|...) \propto \exp\left\{(\boldsymbol{\delta} - \mathbf{b}_1)^\top \mathbf{B}_1^{-1}(\boldsymbol{\delta} - \mathbf{b}_1)\right\}$$

Logo a distribuição condicional a posteriori para $\boldsymbol{\delta}$ é:

$$\boldsymbol{\delta}|... \propto \mathbf{N}_p \left(\begin{array}{l} \left(\mathbf{Z}^\top \mathbf{V}^{-1}\mathbf{Z} + \mathbf{I}_m \otimes \boldsymbol{\Psi}\right)^{-1} \mathbf{Z}^\top \mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k\right), \\ \left(\mathbf{Z}^\top \mathbf{V}^{-1}\mathbf{Z} + \mathbf{I}_m \otimes \boldsymbol{\Psi}\right)^{-1} \end{array} \right) \quad (60)$$

vi) Cálculo da distribuição condicional de $\boldsymbol{\Psi}$

Muitas aplicações estatísticas envolvem vários parâmetros que podem ser considerados como relacionados ou ligados de alguma forma pela estrutura do problema, o que implica que um modelo de probabilidade conjunta para estes parâmetros deve refletir a sua dependência (GELMAN et al., 2013). No nosso caso encontramos esse problema entre a variância específica e sua variância.

Para o cálculo da condicional das variâncias consideramos a distribuição qui-quadrado escalada invertida para cada elemento da matriz que é diagonal com os elementos independentes, ou seja, $\boldsymbol{\Psi} = diag\left(\psi_{11},...,\psi_{kk}\right)$, com $k = 1,..,\mathrm{p}$

Temos:

$$p\left(\delta\right) \propto \left|\boldsymbol{\Psi}^{-1}\right|^{n/2} \exp\left(-\frac{1}{2}\boldsymbol{\delta}^{\top}\boldsymbol{\Psi}^{-1}\boldsymbol{\delta}\right) \tag{61}$$

logo, a distribuição condicional para as variâncias específicaespecíficas é dada como:

$$\propto \prod_{k=1}^{p}\left|\boldsymbol{\Psi}\right|^{-\frac{n+m+2}{2}} \exp\left(-\frac{1}{2}\boldsymbol{\delta}^{\top}\boldsymbol{\Psi}^{-1}\boldsymbol{\delta} + \nu_k S_k^2\boldsymbol{\Psi}^{-1}\right). \tag{62}$$

Reparametrizando (62) e fazendo $S_k^{*2} = \dfrac{\boldsymbol{\delta}^{\top}\boldsymbol{\delta} + \nu_k S_k^2}{m+1}$, $S_k^2 \to 0 \Rightarrow S_k^{*2} = \dfrac{\boldsymbol{\delta}^{\top}\boldsymbol{\delta}}{m+1}$

Fazendo $\eta = m_k + 1$, em que m_k é o número de observações no ambiente,

Mostramos facilmente que a distribuição condicional para cada elemento de $\boldsymbol{\Psi}$ é:

$$\propto \prod_{k=1}^{p}\left(\psi_{kk}\right)^{-\frac{\eta+2}{2}} \exp\left[-\eta S_k^{*2}\right] \tag{63}$$

Que é núcleo de uma qui-quadrado escalonada invertida, apresentada abaixo:

$$\psi_{kk}\,|\,... \sim \mathrm{Inv}-\chi^{-2}\left(m+1, S_k^{*2}\right),\ para\ k = 1,...,\mathrm{p} \tag{64}$$

vii) Distribuição Conjunta a posteriori para $\sigma_{e_k}^2$

Ao considerar como priori a IW para $\mathbf{R}$ a posteriori seria dada por: $p\left(\mathbf{R}\,|\,...\right) \propto N\left[\mathbf{y}\,|\,\mathbf{X}_1\boldsymbol{\beta} + \sum_{k=1}^{p}\lambda_k diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)\mathbf{X}_2\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}, \mathbf{I}_{\mathrm{m}} \otimes \mathbf{R}\right] IW_p\left(\mathbf{R}\,|\,\mathbf{S}_R, \mathrm{p}\right)$

Como vimos $\mathbf{R}\,|\,... \sim IW_p\left(\mathbf{R}\,|\,S_R, \mathrm{p}\right)$, $\mathbf{R} = diag\left(\sigma_{e_1}^2,...,\sigma_{e_k}^2\right)$ que é uma matriz diagonal, não iremos amostrar todos os elementos da matriz diretamente da inversa Whishart. Por conseguinte, foi atribuída como priori a cada elemento da matriz $\mathbf{R}$ uma inversa qui- quadrada escalonada invertida. Dessa a distribuição condicional a posteriori para cada $\sigma_{e_k}^2$ é dada por:

$$P(\sigma_{e_k}^2\,|\,...) \propto \mathrm{N}\left[y_k\,|\,X_{1k}\beta_k + \lambda_k diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)X_{2k}\mathbf{f}_k + Z_k\boldsymbol{\delta}_k, \sigma_k^2\right]\chi^{-2}\left(\sigma_{kk}\,|\,\nu_e, S_{kk}\right)$$

$$\boldsymbol{\varepsilon}_k = \mathbf{y}_k - \left(\mathbf{X}_1\boldsymbol{\beta}_k + \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)\mathbf{X}_2\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta} \right)$$

$$\theta = \left(\mathbf{X}_1\boldsymbol{\beta}_k + \sum_{k=1}^{p} \lambda_k diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)\mathbf{X}_2\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta} \right); \theta_k = X_{1k}\beta_k + \lambda_k diag\left(\mathbf{Z}_1\boldsymbol{\alpha}_k\right)X_{2k}\mathbf{f}_k + Z_k\boldsymbol{\delta}_k$$

$$P(\sigma^2_{e_k} | ...) \propto \left(\sigma^2_{e_k}\right)^{-\frac{m_k}{2}} exp\left\{ -\frac{1}{2\sigma^2_{e_k}} \left(y_k - \theta_k\right)^{\top} \left(y_k - \theta_k\right) \right\} \left(\sigma^2_{e_k}\right)^{-1}$$

$$P(\sigma^2_{e_k} | ...) \propto \left(\sigma_{kk}\right)^{-\left(\frac{m_k}{2}+1\right)} exp\left\{ -\frac{1}{2\sigma^2_k} \left(y_k - \theta_k\right)^{\top} \left(y_k - \theta_k\right) \right\}$$

$$\sigma^2_{e_k} | ... \sim \text{Inv}-\chi^{-2}\left[n_k + 1, \left(y_k - \theta_k\right)^{\top} \left(y_k - \theta_k\right) \right] \quad (65)$$

3.3 Amostragem a partir das distribuições a posteriori condicionais completas dos parâmetros

Como todas as condicionais foram obtidas de forma fechada, ou seja, as distribuições têm formas conhecidas e permitem amostragem direta, o amostrador de Gibbs foi utilizado no processo de estimação dos parâmetros.

Uma observação deve ser feita em relação a amostragem dos vetores singulares. Vimos anteriormente que a condicional para o vetor $\boldsymbol{\alpha}_k$ é uma normal multivariada. Contudo, essa distribuição não satisfaz as restrições do modelo, já que os vetores devem ter norma unitária e serem ortogonais entre si.

A amostragem, entretanto, ainda pode ser realizada a partir da normal multivariada no subespaço corrigido, adicionando dois passos ao algoritmo: normalizar o vetor, e por meio de uma transformação linear devolver os vetores ortogonais no subespaço correto. Viele e Srinivasan (2000) mostraram como realizar a amostragem da uniforme esférica por meio da normal padronizada e como os vetores podem ser colocados no subespaço correto utilizando o processo de ortogonalização Graham Smith.

Vamos assumir que $\boldsymbol{\alpha}_k$ esteja em uma esfera unitária de dimensão m-s no espaço $\Re^m$, ou seja, $\boldsymbol{\alpha}_k$ $\left(\boldsymbol{\alpha}_k \boldsymbol{\alpha}_k = 1\right)$ seja ortogonal a S vetores (*v1, v2,..., vs*) em $\Re^m$. Suponhamos ainda que os vetores $\alpha_1, \alpha_2, ..., \alpha_s$ sejam um conjunto ortonormal de vetores no subespaço gerado por *v1, v2,..., vS*. Nesses termos, dado a matriz $\mathrm{H}_s = \alpha_1, \alpha_2, ..., \alpha_s$, existe uma matriz H_k com as

dimensões $m\times(m-s)$ de forma que $\mathrm{H}_m=(\mathrm{H}_s,\mathrm{H}_k)$ é uma matriz ortonormal. Essa matriz pode ser obtida pelo processo de ortonormalização de Graham Smith.

Assim obtemos o vetor $\tilde{\alpha}_k$ pela transformação linear $\tilde{\alpha}_k=\mathrm{H}_k^{\top}\alpha_k$, de forma que $\tilde{\alpha}_k\in\Re^{m-s}$, ou seja, a amostragem será realizada no subespaço "corrigido" em que a amostragem pode ser conduzida sem as restrições impostas pela decomposição espectral. Pode-se mostrar sem problema que $\alpha_k=\mathrm{H}_k\tilde{\alpha}_k$, ou seja, aplicando a operação inversa recuperamos o vetor no subespaço correto em $\Re^m$ e ortogonais aos *s* vetores, satisfazendo por tanto as restrições do modelo.

A condicional a posteriori obtida em (56) é para um vetor pertencente a $\Re^m$. Assim, é necessário obter a distribuição dos vetores em $\Re^{m-s}$. Para tanto, consideremos a condicional expressa em (56):

$$P(\alpha_k\,|\,...)\propto\exp\left\{-\frac{1}{2}\left(\left[\alpha_k-\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\left[\alpha_k-\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]\right)\right\}$$

Assumindo que $\mathbf{A}_{4k}=\mathbf{X}_1\boldsymbol{\beta}+\sum_{k'\neq k}^{p}\lambda_{k'}diag\left(\mathbf{X}_2\boldsymbol{\alpha}_{k'}\right)\mathbf{Z}_1\mathbf{f}_{k'}+\mathbf{Z}\boldsymbol{\delta}$; $\Delta_k=\lambda_k diag\left(\mathbf{Z}_1\mathbf{f}_k\right)\mathbf{X}_2$ e resolvendo dentro dos colchetes temos:

$$\alpha_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\alpha_k-\alpha_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\left(\Delta^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}-\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\alpha_k$$
$$+\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}$$

Utilizando agora a identidade $\mathrm{H}_k\mathrm{H}_k^{\top}=\mathbf{I}$ e desenvolvendo os cálculos têm-se as seguintes parcelas da soma acima:

i) $$\alpha_k^{\top}\mathrm{H}_k\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_r^{\top}\alpha_k=(\mathrm{H}_k^{\top}\alpha_k)^{\top}\mathrm{H}_r^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^{\top}\alpha_k=$$
$$=(\tilde{\alpha}_k)^{\top}\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\tilde{\alpha}_k$$

ii) $$-\alpha_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}=-\alpha_k^{\top}\mathrm{H}_k\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}$$
$$=-\tilde{\alpha}_k^{\top}\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}$$

iii) $$-\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^{\top}\alpha_k=-\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k$$
$$=-\left[\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^{\top}\mathbf{R}^{-1}\mathbf{A}_{4k}\right]^{\top}\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\tilde{\alpha}_k$$

$$\text{iv)}\left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k} =$$

$$=\left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}$$

Dessa forma:

$$\alpha_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\alpha_k - \alpha_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k} - \left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\alpha_k$$
$$+\left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k} =$$

$$(\tilde{\alpha}_k)^\top \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\tilde{\alpha}_k - \tilde{\alpha}_k^\top \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k} -$$
$$\left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)H\tilde{\alpha}_k +$$
$$\left[\left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\right]^\top \mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}$$

Assim, a condicional a posteriori para a variável auxiliar $\tilde{\alpha}_k$, dados os demais parâmetros, no subespaço corrigido é dada por:

$$P(\tilde{\boldsymbol{\alpha}}_k \mid ...) \propto \exp\left\{-\frac{1}{2}\left(\begin{bmatrix}\tilde{\alpha}_k - \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\end{bmatrix}^\top \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k \times \begin{bmatrix}\tilde{\alpha}_k - \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}\end{bmatrix}\right)\right\}$$

$$\tilde{\boldsymbol{\alpha}}_k \mid ... \sim N\left[\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}, \left(\mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k\right)^{-1}\right].$$

Dessa forma, amostragem de $\tilde{\alpha}_k$ é realizada no subespaço corrigido por meio da condicional obtida anteriormente. Assim, como apresentando em Crossa et al. (2011) e Oliveira et al. (2015), o interesse é amostrar vetores que tenham norma 1, logo após esse primeiro passo o vetor deve ser normalizado, $\boldsymbol{\alpha}_k^* = \dfrac{\tilde{\alpha}_k}{\sqrt{\tilde{\alpha}_r^\top \tilde{\alpha}_k}}$ e assumindo $\bar{\tilde{\alpha}}_k = \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)^{-1}\Delta_k^\top \mathbf{R}^{-1}\mathbf{A}_{4k}$ e $C_k = \sqrt{\tilde{\alpha}_r^\top \tilde{\alpha}_k}$, por meio de manipulações algébricas chegamos a condicional para os $\tilde{\alpha}_k$ com norma 1 e ortogonais entre si dada por:

$$\tilde{\boldsymbol{\alpha}}_k \mid ... \sim N\left[\boldsymbol{\alpha}_k^*, \left(\mathrm{C}_k^\top \mathrm{H}_k^\top \left(\Delta_k^\top \mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k C_k\right)^{-1}\right]$$

em que que: $\boldsymbol{\alpha}_k^* = \dfrac{\tilde{\alpha}_k}{C_k}$.

Para retornar o vetor singular ao correto subespaço em $\Re^m$, satisfazendo as restrições de ortonormalidade, basta aplicar a transformação inversa $\alpha_k = \mathrm{H}_k \tilde{\alpha}_k$. O vetor é ortogonal aos outros *s* vetores e a transformação preserva a norma do vetor, pois

$$(\tilde{\alpha}_k)^\top \tilde{\alpha}_k = (\mathrm{H}_k^\top \alpha_k)^\top \mathrm{H}_k^\top \alpha_k = \alpha_k^\top \mathrm{H}_R \mathrm{H}_R^\top \alpha_k = \alpha_k^\top \mathrm{I}_m \alpha_k^\top = \alpha_k^\top \alpha_k = 1.$$

dessa forma, um vetor aleatório $m-s$ dimensional em $\Re^m$ pode ser transformado um a um no mesmo vetor aleatório em $\Re^{m-s}$, como foi demonstrado por Liu (2001) para o caso da distribuição Von Mises Fisher.

3.3.1 Algoritmo para o amostrador de Gibbs

A partir das distribuições condicionais completas, como obtidas anteriormente para cada parâmetro ou conjunto de parâmetros foram obtidas cadeias de Markov, pelo método de Monte Carlo (MCMC). Lembrando que todas as condicionais permitem amostragem direta, utilizou-se apenas o amostrador de Gibbs.

O algoritmo iterativo de amostragem implementado pode ser ilustrado por meio dos seguintes passos:

1 - Primeiramente são atribuídos valores iniciais aos parâmetros do modelo:

$$\Theta_1^0 = \left[\boldsymbol{\beta}^0, \mathbf{R}^0, \boldsymbol{\lambda}^0, \mathbf{f}^0, \boldsymbol{\alpha}^0, \boldsymbol{\delta}^0, \boldsymbol{\Psi}^0\right]$$

2- A partir desses valores iniciais, a i -ésima iteração pode ser obtida da seguinte forma:

a) Gerar $\boldsymbol{\beta}^i \mid \mathbf{R}^{i-1}, \lambda^{i-1}, \boldsymbol{\alpha}^{i-1}, \mathbf{f}^{i-1}, \boldsymbol{\delta}^{i-1}, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori

$$\boldsymbol{\beta} \mid ... \alpha N\left(\left(\mathbf{X}_1^\top \mathbf{R}^{-1} \mathbf{X}_1\right)^{-1} \mathbf{X}_1^\top \mathbf{R}^{-1} \left\{\mathbf{y} - \left(\sum_{k=1}^{p} \lambda_k diag\left(\mathbf{X}_2 \boldsymbol{\alpha}_k\right) \mathbf{Z}_1 \mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}\right)\right\}, \left(\mathbf{X}_1^\top \mathbf{R}^{-1} \mathbf{X}_1\right)^{-1}\right)$$

b) Gerar $\left(\sigma_{e_k}^2\right)^{i-1} \mid \boldsymbol{\beta}^i, \boldsymbol{\lambda}^{i-1}, \boldsymbol{\alpha}^{i-1}, \mathbf{f}^{i-1}, \boldsymbol{\delta}^{i-1}, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori

$$\sigma_{e_k}^2 \mid ... \propto \mathrm{Inv}\text{-}\chi^{-2}\left[n_k + 1, \left(y_k - \theta_k\right)^\top \left(y_k - \theta_k\right)\right]$$

c) Para gerar a i-ésima observação dos valores singulares e vetores singulares devem ser seguida a sequência c1), e c2), abaixo, para k= 1, 2, ..., p

c1) Gerar $\lambda_k^i \,|\, \boldsymbol{\beta}^i, \mathbf{R}^i, \alpha_k^{i-1}, f_k^{i-1}, \boldsymbol{\delta}^{i-1}, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori

$$\lambda_k \,|\, ... \propto N^{+}\left[\left(\mathbf{A}_{2k}^{\top}\mathbf{R}^{-1}\mathbf{A}_{2k}\right)^{-1}\mathbf{A}_{2k}^{\top}\mathbf{R}^{-1}\mathbf{A}_1, \left(\mathbf{A}_{2k}^{\top}\mathbf{R}^{-1}\mathbf{A}_{2k}\right)^{-1}\right],$$

com: $A_{1k} = \mathbf{y} - \left[\mathbf{X}_1\beta + \sum_{k=1}^{m}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k + \mathbf{Z}\boldsymbol{\delta}\right]$ e $A_{2k} = diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k$

c2) Gerar $\boldsymbol{\alpha}_k^i \,|\, \boldsymbol{\beta}^i, \mathbf{R}^i, \lambda_k^i, f_k^{i-1}, \boldsymbol{\delta}^{i-1}, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori.

i) Amostrar $\tilde{\boldsymbol{\alpha}}_k$ a partir da distribuição

$$\tilde{\boldsymbol{\alpha}}_k \,|\, ... \sim N\left[\boldsymbol{\alpha}_k^*, \left(\mathrm{C}_k^{\top}\mathrm{H}_k^{\top}\left(\Delta_k^{\top}\mathbf{R}^{-1}\Delta_k\right)\mathrm{H}_k C_k\right)^{-1}\right]$$

ii) Aplicar a transformação $\alpha_k = \mathrm{H}_r\alpha_k^*$ obtendo o vetor no correto subespaço em R^{m} .

d) Gerar $\mathbf{f} \,|\, \boldsymbol{\beta}^i, \mathbf{R}^i, \boldsymbol{\alpha}^i, \boldsymbol{\lambda}^i, \boldsymbol{\delta}^{i-1}, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori

$$\mathbf{f} \,|\, ... \sim \mathbf{N}_p\left(\left(\mathbf{I} + \mathbf{A}_3^{\top}\mathbf{R}^{-1}\mathbf{A}_3\right)^{-1}\mathbf{A}_3^{\top}\mathbf{R}^{-1}\varphi; ; \left(\mathbf{I} + \mathbf{A}_3^{\top}\mathbf{R}^{-1}\mathbf{A}_3\right)^{-1}\right)$$

em que $\varphi = \mathbf{y} - \mathbf{X}_1\beta - \mathbf{Z}\boldsymbol{\delta}$; $\mathbf{A}_3 = \left(\sum_{k=1}^{m}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\right)$

e) Gerar $\boldsymbol{\delta}^i \,|\, \boldsymbol{\beta}^i, \mathbf{R}^i, \boldsymbol{\alpha}^i, \lambda^i, \mathbf{f}^i, \boldsymbol{\Psi}^{i-1}$ a partir da distribuição condicional a posteriori

$$\boldsymbol{\delta} \,|\, ... \propto \mathbf{N}_p\left(\left(\mathbf{Z}^{\top}\mathbf{V}^{-1}\mathbf{Z} + \mathbf{I}_{\mathbf{p}} \otimes \boldsymbol{\Psi}\right)^{-1}\mathbf{Z}^{\top}\mathbf{R}^{-1}\left(\mathbf{y} - \mathbf{X}_1\boldsymbol{\beta} - \sum_{k=1}^{m}\lambda_k diag\left(\mathbf{X}_2\boldsymbol{\alpha}_k\right)\mathbf{Z}_1\mathbf{f}_k\right), \left(\mathbf{Z}^{\top}\mathbf{V}^{-1}\mathbf{Z} + \mathbf{I}_{\mathbf{p}} \otimes \boldsymbol{\Psi}\right)^{-1}\right)$$

f) Gerar $\boldsymbol{\Psi}^i \,|\, \boldsymbol{\beta}^i, \mathbf{R}^i, \lambda^i\boldsymbol{\alpha}^i, \mathbf{f}^i, \boldsymbol{\delta}^i$ a partir da distribuição condicional a posteriori

$$\psi_{kk} \,|\, ... \sim \mathrm{Inv}-\chi^{-2}\left(m+1, S_k^{*2}\right), \; para \; k = 1, ..., \mathrm{p}\,.$$

A implementação???? do processo de amostragem de Gibbs está baseada na geração de cadeias de Markov e amostras sucessivas de “a” até “f” , descartando-se um conjunto inicial de amostras para garantir a convergência da cadeia para a sua distribuição estacionária e para eliminar o efeito dos valores iniciais. Essas amostras após o período de descarte são representativas das densidades posterioris marginais dos parâmetros do modelo, das quais é efetuada a inferência sobre esses parâmetros ou quaisquer funções de interesse envolvendo tais parâmetros.

Constatando a convergência das cadeias, as amostras obtidas são usadas na estimação dos parâmetros e construção de intervalos de credibilidade. O diagnóstico de convergência das cadeias produzidas, de acordo com os passos descritos acima, foi realizado pelo método de

Raftery e Lewis (1992) e pelo critério de Heidelberger e Welch (1983). Todo processo de inferência foi realizado utilizando o software estatístico R (R CORE TEAM, 2016).

3.4. Inferência sobre os parâmetros lineares e multiplicativos do modelo.

As estimativas para $\boldsymbol{\beta}$, $\mathbf{f}$, τ_{ji}, $\mathbf{G}\left(\sigma^2_{ge_{ii}}, \sigma^2_{ge_{ij}}\right)$, $\mathbf{R}\left(\sigma^2_{e_1}, ..., \sigma^2_{e_k}\right)$, e $\boldsymbol{\Psi}$ foram obtidas pelas médias das amostras MCMC. No caso dos vetores singulares as médias amostrais não satisfazem as restrições do modelo, no que se refere à ortonormalidade. As regiões de máxima credibilidade a posteriori para os parâmetros serão construídas de acordo com o método proposto por Chen e Shao (1999) implementado no pacote BOA (*bayesian output analysis*) utilizando o software estatístico R (R CORE TEAM, 2016).

3.4.1 Regiões de credibilidade bivariadas para os escores fatoriais genotípicos e ambientais no biplot

As regiões de credibilidade no biplot serão construídas para as duas primeiras cargas fatoriais, em relação aos escores $(\tau_{j1}\mathbf{f}_{i1},\ \tau_{j2}\mathbf{f}_{i2})$ e com $i = 1, \ldots, \text{m}$ e $\text{j} = 1, \ldots, \text{p}$, para genótipos e ambientes, respectivamente, usando a expressão $\mathbf{u}_{g_{ij}} = \boldsymbol{\tau}_{j1}\mathbf{f}_{i1} + \boldsymbol{\tau}_{j2}\mathbf{f}_{i2} + \boldsymbol{\delta}_{ij}$, em que $\mathbf{f}$ e $\boldsymbol{\delta}$ são como definidos anteriormente. O método aqui utilizado para encontrar regiões de credibilidade, para cada par de escores genotípicos e ambientais, foi calcular as distâncias euclidianas dos pontos amostrais, em relação ao centro da distribuição, para cada observação e remover os 5% de pontos com a maior distância (OOMS, 2009).

Primeiramente são obtidas as versões padronizadas das duas variáveis que compõem a amostra bivariada. Para cada variável, a padronização é realizada subtraindo a respectiva média para cada observação e dividindo o resultado pelo desvio padrão. A partir da distribuição bivariada padronizada (*d*), calcula-se a distância euclidiana de cada ponto ao centro da distribuição, representado pelo ponto contendo as médias das variáveis. Feito isso, é calculado o quantil 95% para a distribuição dessas distâncias, que pode ser denotado por *r*. Em seguida, todos os pontos com uma distância euclidiana, em relação ao centro, maior que *r* são removidas da amostra.

O último passo é transformar os pontos amostrais para o dimensionamento da distribuição original. Para tanto, as observações de cada variável, que compõe a distribuição bivariada, são multiplicadas pelos respectivos desvios padrão e as respectivas médias devem ser adicionadas a estes valores. Com este procedimento são obtidas regiões de credibilidade com formas elípticas, que abrangem os 95% de pontos da distribuição original. Este procedimento pode ser realizado para obter outros níveis de credibilidade.

Sobreposições entre as regiões de credibilidade indicam que os valores dos escores não são estatisticamente diferentes entre si e, portanto, os respectivos genótipos ou ambientes fazem parte de um subgrupo homogêneo, ou seja, com características semelhantes em relação à interação GEI. Por outro lado, valores da interação, referentes aos escores cujas regiões de credibilidade englobem a origem (0,0), são considerados não estatisticamente diferentes de zero, no nível , ou , em nível de...a expressão (ao nível) é equivocada de credibilidade considerado. Os genótipos (ou ambientes) relacionados aos escores cujas regiões de credibilidade abrangem a origem são considerados estáveis, ou seja, não contribuem significativamente para a interação GEI tal qual ressaltado por Crossa et al. (2011) e Oliveira et al. (2015).

3.5. Validação do modelo na predição de dados faltantes

Assumindo mesmas prioris e posterioris e processo de inferência para os dados balanceados, o processo será repetido com dados desbalanceados. O desbalanceamento nos dados será feito de forma aleatória dentro de cada ambiente (consistindo em perda total do genótipo), e separando a população de treino e de validação, considerando níveis de perda de 10%, 33% e 50%.

Embora não fosse o propósito do trabalho, para validação do modelo foi feita a comparação com o modelo FA padrão, em dois estágios apresentado em Nuvunga et al. (2015) via algoritmo EM (FA-EM-expectation maximization) implementado em R e SAS e em um estágio de Smith, Cullis e Thompson (2001), via algoritmo AI (FA-AI-average information) em matrizes esparsas implementado???? em Asreml-R (BUTLER et al., 2009).

A capacidade preditiva dos modelos será medida calculando-se a PRESS-média *(Predicted Residual Sum of Squares***)** **e** a correlação entre o valor fenotípico predito ($\hat{y}_{ij}$) e observado (y_{ij}).

A expressão da PRESS é dada por:

$$PRESS = \frac{1}{n}\sum_{j=1}^{n}\left(y_{ij} - \hat{y}_{ij}\right)^2 \qquad (66)$$

e da correlação:

$$correlação = \frac{\sum_{i=1}^{n}\left(\hat{y}_{ij} - \bar{\hat{y}}_{ij}\right)\left(y_{ij} - \bar{y}_{ij}\right)}{\sqrt{\sum_{i=1}^{n}\left(\hat{y}_{ij} - \bar{\hat{y}}_{ij}\right)^2}\sqrt{\sum_{i=1}^{n}\left(y_{ij} - \bar{y}_{ij}\right)^2}} \quad . \qquad (67)$$

3.6. Seleção do modelo ou escolha do número de fatores k

Uma vez que foram usados dados simulados e reais, a seleção de modelos que em análise de fatores significa escolha no número de cargas ou fatores k que explicam que os dados foram adotados critérios diferentes para cada conjunto de dados, um baseado em PRESS e o outro em Akaike Monte Carlo, descritos a seguir.

3.6.1 Dados simulados

A seleção de modelos para este conjunto de dados foi feita com base no critério PRESS descrito por Gabriel (2002) que usa abordagem de validação cruzada e o critério de eficiência estatística dada pela $SE = PRESS - full / PRESS - k$ que refere a razão entre a PRESS do modelo com todos componentes contra PRESS- relacionados ao fator de ordem k selecionando o modelo que apresentar maior SE e menor PRESS.

3.6.2. Dados experimentais

Para comparar diferentes modelos foram utilizadas técnicas de estimação de verossimilhanças marginais via critério de informação de Akaike Monte Carlo (AICM) apresentados por Raftery et al. (2007) e $\Delta AICM$ que corresponde a diferença entre o modelo completo e os demais modelos em competição.

O AICM é calculado a partir da saída MCMC da análise bayesiana.

O AICM é calculado como versão baseada em simulação posterior da AIC (AKAIKE, 1973):

$$AICM = 2\left(\overline{l} - s_l^2\right) \qquad (68)$$

em que: $\overline{l}$ - é a média a posteriori da log verossimilhança marginal , s_l^2-é a variância a posteriori da log verossimilhança marginal.

Assim, o AICM é visto como uma versão muito simples do cálculo penalizado da média posteriori do log verossimilhança, utilizando apenas a log verossimilhança da simulação posteriori. Sendo preferido o modelo que apresentar o maior AICM e menor $\Delta AICM$.

4. RESULTADOS E DISCUSSSÃO

Nesta seção, objetiva-se apresentar os principais resultados obtidos a partir do método proposto e discussão dos resultados mais importantes.

4.1. RESULTADOS

4.1.1. Conjunto de dados simulados

Inicialmente foi realizado um estudo de simulação para ilustrar o desempenho da versão Bayesiana do modelo fatorial analítico (FAB) proposto. Foram simulados diferentes valores (os chamados valores paramétricos) para comparar com as estimativas a posteriori do modelo proposto. O conjunto de dados simulados é composto por 20 genótipos e cinco ambientes.

Os valores verdadeiros dos parâmetros dos efeitos fixos e variância residual usados na simulação para gerar os dados foram: $\boldsymbol{\beta} = (9,980; 11,109; 10,817; 12,014; 10,711)$ e $\sigma^2_{e_k} = (0,546; 1,209; 4,690; 7,377; 9,026)$, e as médias geométricas da variância genética e residual (raiz quinta do produto das variâncias de cada ambiente) foram 5,90 e 2,90 respectivamente. A matriz de cargas proveniente da decomposição espectral da matriz de covariâncias genética é dada por:

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1,468 & 0,579 & -0,658 & 0,090 & -0,586 \\ 1,314 & 0,385 & -0,347 & 1,034 & 0,194 \\ 2,404 & 0,257 & -1,275 & -0,503 & 0,273 \\ 1,771 & 1,717 & 1,320 & -0,176 & 0,055 \\ 2,751 & -1,823 & 0,781 & 0,010 & -0,054 \end{bmatrix}$$

Portanto, a matriz de covariâncias genética em cada ambiente é dada por:

$$\boldsymbol{\Sigma} = \begin{bmatrix} 3,273 & & & & \\ 2,359 & 3,103 & & & \\ 4,311 & 3,239 & 7,800 & & \\ 2,676 & 2,360 & 3,122 & 7,861 & \\ 2,501 & 2,644 & 5,130 & 2,770 & 11,505 \end{bmatrix}$$

Para análise bayesiana, foram simuladas cadeias MCMC com 58 mil iterações para os parâmetros do modelo FAB. As primeiras 840 observações foram descartadas e as amostras foram obtidas realizando-se saltos a cada 4 observações, no intuito de evitar amostragem de cadeias que ainda não tivessem atingido convergência e não selecionar observações correlacionadas. Ao final desse processo foram obtidas amostras com 14290 observações, para cada parâmetro.

Como especificado na seção Métodos, a convergência das cadeias de Markov, geradas pelo amostrador de Gibbs, foi monitorada utilizando os critérios de Raftery e Lewis (1992), teste de Heidelberger e Welch e teste de Gweke (GELMAN et al., 2013). Esses testes indicaram boas propriedades de convergência para todos os parâmetros do modelo. Nas Figuras (16-17 em apêndice-B) apresentam-se os traços das cadeias geradas para os valores da variância residual, ficando nítido o agrupamento dos valores, em todas as cadeias, em torno de um valor especifico, corroborando, portanto, os resultados dos testes de convergência. Fato semelhante foi observado para os demais parâmetros.

Os valores paramêtricos e médias a posteriori para os efeitos lineares, variância residual e dos efeitos ambientais (valores singulares e vetores singulares) são apresentados na Tabela 1. Comparações dos valores dos desvios entre os valores estimados pelo modelo e os valores paramétricos foram feitos. O modelo funcionou muito bem, com estimativas (médias a posteriori) muito perto dos valores verdadeiros usados para simular os dados.

Tabela 1- Médias a posteriori (MP), valor paramétrico (VD), desvio padrão (sd), intervalos de credibilidade (I.C. 95%, LI: limite inferior, LS: limite superior) e os valores verdadeiros para os efeitos ambientais ($\beta_1 - \beta_5$), variância residual ($\sigma^2_{e_1} - \sigma^2_{e_5}$), para as duas primeiras cargas fatoriais (efeitos ambientais) ($\tau_{j1} - \tau_{j2}$) .(continua)

				Intervalo HPD 95%	
Parâmetro	VD	MP	sd	LI	LS
β_1	9,980	10,625	0,462	9,744	11,564
β_2	11,109	11,664	0,449	10,756	12,5478
β_3	10,817	11,749	0,716	10,369	13,183
β_4	12,014	12,821	0,688	11,475	14,1804
β_5	10,711	12,044	0,794	10,532	13,6739
$\sigma^2_{e_1}$	0,546	0,8471	1,260	0,299	1,5064
$\sigma^2_{e_2}$	1,209	1,911	1,155	0,846	3,1618
$\sigma^2_{e_3}$	4,690	4,381	2,353	2,209	6,826
$\sigma^2_{e_4}$	7,377	10,091	19,936	4,938	15,097
$\sigma^2_{e_5}$	9,026	9,104	3,807	3,503	15,7934
σ^2_e	2,900	2,848	-	-	-
τ_{11}	1,468	1,548	0,845	0,500	3,448
τ_{21}	1,314	1,344	0,761	0,291	3,062

Tabela 2- Médias a posteriori (MP), valor paramétrico (VD), desvio padrão (sd), regiões de credibilidade (I.C. 95%, LI: limite inferior, LS: limite superior) e os valores verdadeiros para os efeitos ambientais ($\beta_1 - \beta_5$), variância residual ($\sigma^2_{e_1} - \sigma^2_{e_5}$), para as duas primeiras cargas fatoriais (efeitos ambientais) ($\tau_{j1} - \tau_{j2}$) .(conclusão)

				Intervalo HPD 95%	
Parâmetro	VD	MP	sd	LI	LS
τ_{31}	2,404	2,353	1,267	0,799	4,916
τ_{41}	1,771	1,475	1,008	-0,439	3,817
τ_{12}	0,579	0,381	0,603	-0,845	1,521
τ_{22}	0,385	0,436	0,374	0,000	1,154
τ_{32}	0,257	0,234	0,678	-1,093	1,657
τ_{42}	1,717	0,307	0,890	-1,466	2,113
τ_{52}	-1,823	-0,764	1,154	-2,726	1,748

Os valores dos efeitos ambientais e das variâncias residuais estimados foram próximos dos valores verdadeiros utilizados para simulação. Observou-se que todos valores utiolizados nas simulações estão dentro dos intervalos de credibilidade considerados. As coordenadas das cargas fatoriais relacionadas aos dois primeiros eixos para os ambientes foram praticamente iguais aos valores verdadeiros e dentro do raio de cobertura dos intervalos de credibilidade a 95% de probabilidade.

Estimativas pontuais e regiões de credibilidade para as coordenadas relacionadas aos dois primeiros eixos, para genótipos, podem ser observadas na Tabela 2. É possível perceber que as estimativas a posteriori apresentam valores muito próximos às estimativas de máxima verossimilhança restrita (REML) via algoritmo average information (FA-AI) das coordenadas dos escores genotípicos, referentes ao primeiro eixo. Para o segundo eixo, percebe-se maior diferença entre as médias a posteriori e as estimativas REML do FA-AI para o G1, G2 e G5 com seus valores dentro dos intervalos de credibilidade.

Tabela 3- Médias a posteriori (MP), regiões de credibilidade (I.C. 95%, LI: limite inferior, LS: limite superior) e estimativas de máxima verossimilhança restrita (REML) do FA-AI e escores genotípicos ($\mathbf{f}_{i1} - \mathbf{f}_{i2}$).

Par.	REML	MP	sd	LI	LI	Par.	REML	MP	sd	LI	LS
f_{11}	-0,959	-1,141	0,423	-1,970	-0,310	f_{12}	-0,529	0,062	0,904	-1,722	1,837
f_{21}	-1,351	-1,261	0,445	-2,135	-0,414	f_{22}	-0,024	0,454	0,908	-1,447	2,173
f_{31}	0,043	-0,093	0,367	-0,824	0,621	f_{32}	-0,463	-0,247	0,733	-1,699	1,238
f_{41}	1,127	0,496	0,481	-0,456	1,410	f_{42}	-1,538	-1,092	0,966	-2,778	1,009
f_{51}	-0,841	-0,787	0,393	-1,607	-0,064	f_{52}	-0,134	0,203	0,811	-1,357	1,846
f_{61}	0,735	1,066	0,452	0,222	1,970	f_{62}	1,438	0,581	0,809	-0,992	2,200
f_{71}	-0,294	-0,222	0,375	-0,956	0,512	f_{72}	-0,206	-0,115	0,819	-1,778	1,479
f_{81}	-1,229	-0,816	0,425	-1,684	-0,008	f_{82}	0,692	0,659	0,871	-1,183	2,266
f_{91}	-1,678	-1,123	0,474	-2,056	-0,211	f_{92}	0,762	0,712	1,081	-1,495	2,741
f_{101}	-0,526	-0,601	0,386	-1,397	0,104	f_{102}	-0,627	-0,247	0,794	-1,844	1,345
f_{111}	-0,149	-0,471	0,397	-1,281	0,277	f_{112}	-0,965	-0,433	0,788	-1,975	1,163
f_{121}	-0,172	-0,146	0,367	-0,871	0,577	f_{122}	-0,046	-0,038	0,757	-1,565	1,475
f_{131}	0,509	0,900	0,454	0,046	1,796	f_{132}	1,566	0,717	0,880	-1,087	2,413
f_{141}	1,270	1,156	0,419	0,371	2,002	f_{142}	-0,085	-0,421	0,834	-2,029	1,249
f_{151}	0,078	0,239	0,378	-0,492	0,984	f_{152}	0,617	0,312	0,768	-1,233	1,845
f_{161}	1,439	1,498	0,444	0,642	2,387	f_{162}	0,655	-0,044	0,856	-1,678	1,669
f_{171}	0,745	0,565	0,399	-0,2192	1,356	f_{172}	-0,538	-0,519	0,804	-2,006	1,186
f_{181}	1,369	1,424	0,446	0,555	2,291	f_{182}	0,938	0,150	0,847	-1,528	1,802
f_{191}	0,825	0,359	0,413	-0,484	1,144	f_{192}	-0,934	-0,653	0,831	-2,259	1,041
f_{201}	-0,936	-1,024	0,401	-1,848	-0,275	f_{202}	-0,589	-0,058	0,780	-1,625	1,446

Par= Parâmetro

Em geral, as estimativas dos parâmetros estão dentro dos intervalos estimados e próximo dos verdadeiros valores. Para facilitar a compreensão, apenas os dados relevantes obtidos a partir da análise dos dados foram representados, uma vez que valores correspondentes para ambientes individuais foram semelhantes. As regiões bivariadas (HPD) das cargas fatorais (τ_{j1}, τ_{j2}) e escores genotípicos $(\mathbf{f}_{i1}, \mathbf{f}_{i2})$ que foram estatisticamente diferentes da origem (0,0) estão representados nas Figuras 1-2.

A Figura 1 mostra as regiões HPD para os ambientes considerando (τ_{j1}, τ_{j2}) que são representados pelas cargas fatoriais que por definição são predições das médias de genótipos para um ambiente que é "médio" no sentido de ter covariância média (em termos de efeitos de genótipos) com todos os outros ambientes. Portanto, em modelos fatores analíticos sem efeitos principais de G, as cargas são úteis para agrupamento ambientes em termos de correlações genéticas (SMITH; CULLIS; THOMPSON, 2001). Vale ressaltar neste modelo os efeitos de G confundidos com a GEI, isto é, sem efeitos principais de genótipos, logo a primeira carga deve ter sinal positivo. Segundo Burgueño et al. (2008), Smith, Cullis e Thompson (2001) e Stefanova e Buirchell (2010), o primeiro fator explica a quantidade máxima do efeito interação GEI nos dados, e assim por diante. Dessa forma a interpretação do biplot se os escores genotípicos e ambientais fossem representados no mesmo gráfico seria similar à do modelo GGE-biplot, em que o primeiro componente principal captura boa parte do efeito principal de genótipos e o segundo a interação GE. Podemos observar que a escala de valores foi equivalente entre os E-BLUPS e o primeiro escore fatorial dado o ajuste da regressão que foi de praticamente 1 (r^2=0,979).

É notável a formação de um cluster entre os ambientes {E1, E2, E4}, que são os ambientes mais correlacionados, e os ambientes {E3, E5} estão menos correlacionados entre si. Porém como temos uma clara sobreposição entre os seus HPD a 95% de probabilidade estes grupos não seriam estatisticamente diferentes entre si.

A Figura 2 mostra as regiões bivariadas dos escores genotípicos $(\mathbf{f}_{i1},\mathbf{f}_{i2})$, a 95% de probabilidade. Podemos verificar a formação de dois subgrupos de genótipos distintos entre si, que é composto pelos genótipos {G1, G2, G9, G20} e {G6, G14, G16, G18}, que se encontram no mesmo quadrante e dentro de cada subgrupo os genótipos não são distintos entre si, ou seja, apresentam efeitos semelhantes com relação a interação, a 95% de probabilidade no HPD.

Ao longo dos ambientes, o modelo proposto permite inferir que os genótipos G16 e G18 tenham a melhor classificação nos ambientes (mesma direção, mesmo quadrante). Os genótipos G1, G2, G9 e G20 estão localizados na direção oposta no biplot, e seriam mal classificados nos ambientes de teste.

Embora a interpretação do modelo seja similar ao GGE-biplot essa deveria ser vista com cautela, pois os nossos gráficos foram representados separados e no segundo eixo-FA2, não é possível verificar a separação de grupos de genótipos ou ambientes, bem como não é possível fazer o produto interno. No caso de genótipos, como as regiões cortam o primeiro eixo, todos seriam considerados estáveis.

Figura 1-Regiões de credibilidade a 95% de probabilidade para a cargas fatoriais dos ambientes cujas regiões não englobam a origem.

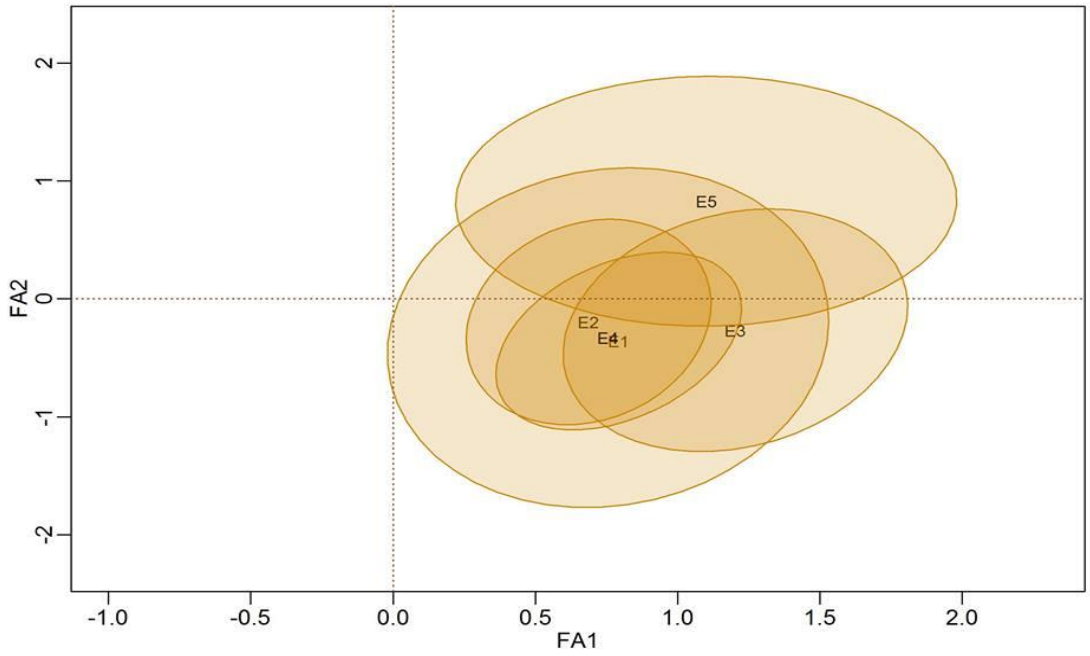
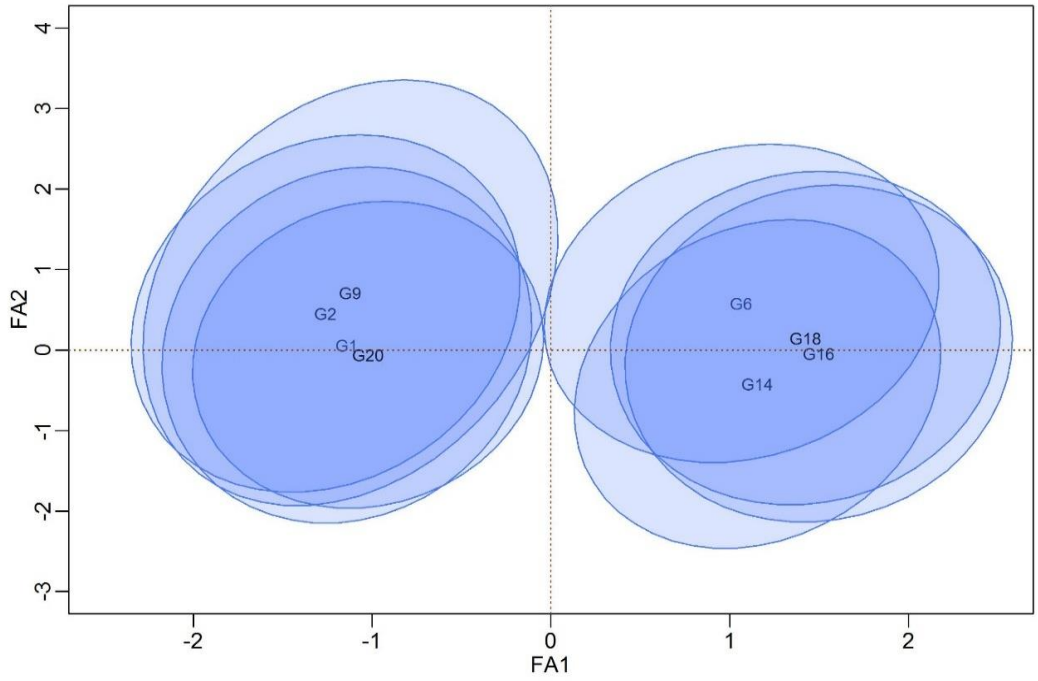


Figura 2-Regiões de credibilidade a 95% de probabilidade para os escores fatoriais genotípicos que não englobam a origem.



4.1.2. Avaliação do desempenho do modelo sob desbalanceamento

Para verificar a robustez do modelo bayesiano proposto na predição dos dados faltantes realizou-se a comparação com a versão mista do modelo FA via algoritmo EM (FA-EM) ajustado em dois estágios (Nuvunga et al. 2015), e FA-AI em um estágio (BUTLER et al., 2009), considerando modelos completos. Embora seja difícil ajustar um modelo completo em modelos FA baseados em modelos mistos devido ao custo computacional e problemas de convergência, esse problema não foi observado, pois o conjunto de ambientes é baixo.

Foram simuladas perdas aleatórias (completa) dos genótipos nos ambientes de 10%, 33% e 50%.

Os resultados de validação cruzada demonstraram que é possível predizer o desempenho de híbridos utilizando modelos FA com elevada acurácia (0,82). Nas figuras 3-5, verifica-se que a correlação foi de moderada magnitude (0,18-0,86) para todos os níveis de desbalanceamento, com magnitude inversamente proporcional ao nível de perda utilizado. É possível verificar que o modelo bayesiano tem uma capacidade preditiva maior que o modelo FA misto (FA-EM e FA-AI) (mesmo usando o critério PRESS para seleção o modelo FA seria FAB seria o preferido).

Dado que as perdas dos genótipos foram aleatórias, a porcentagem de híbridos desbalanceados no conjunto de dados também variou a cada ciclo. Por exemplo, considerando um nível de 10% de perdas de genótipo, o número total de genótipos perdidos variou de 1 a 6 por ambiente, porém verifica-se uma clara diferença nas correlações desses resultados (Fig. 3 a 5) que podem estar relacionados ao tipo de genótipo perdido (não adaptado/estável).

Figura 3-Gráfico de barras da correlação proveniente da validação cruzada considerando desbalanceamento de 10% de perda de genótipos nos ambientes usando o modelo FA bayesiano (FAB) FA padrão (FA-EM e FA-AI).

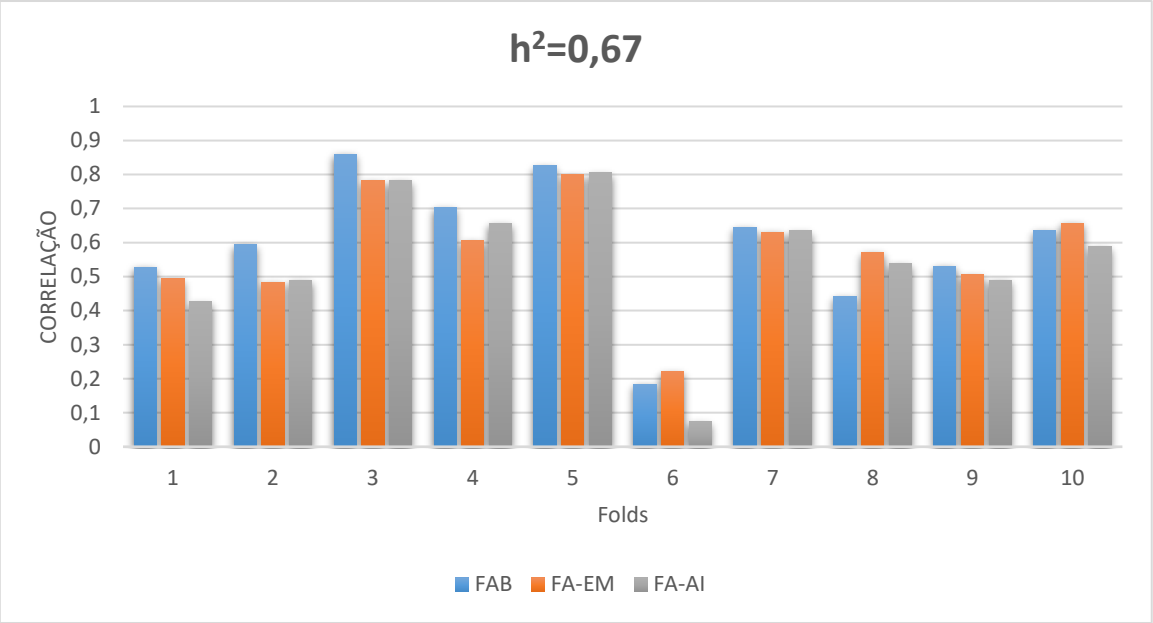


Figura 4-Gráfico de barras da correlação proveniente da validação cruzada considerando desbalanceamento de 33% de perda de genótipos nos ambientes usando o modelo FA bayesiano (FAB) FA padrão (FA-EM e FA-AI).

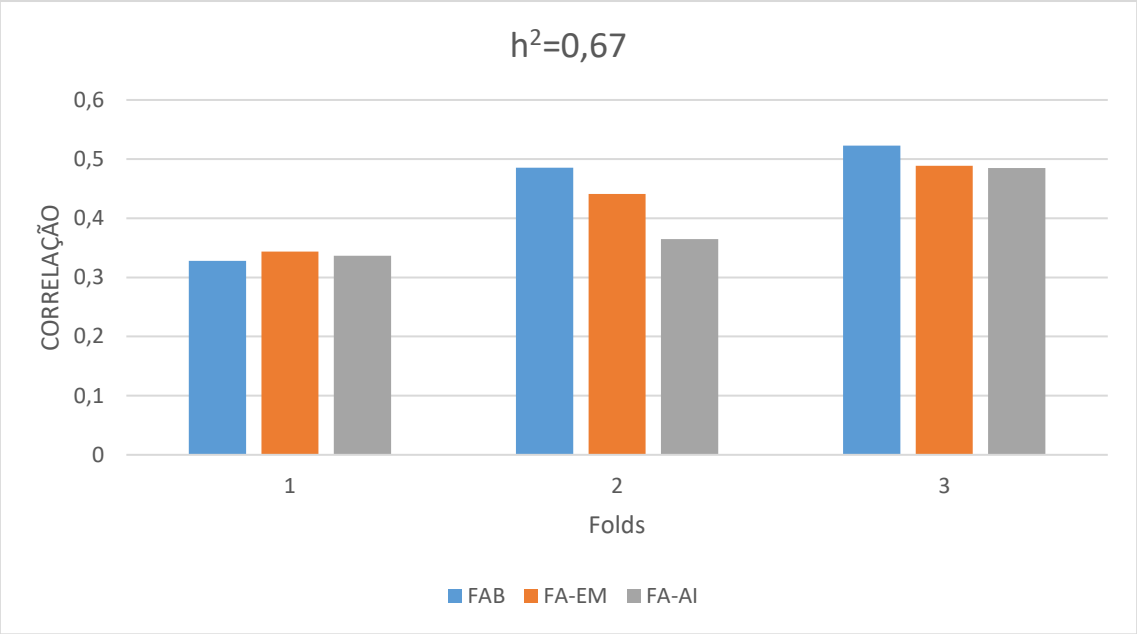
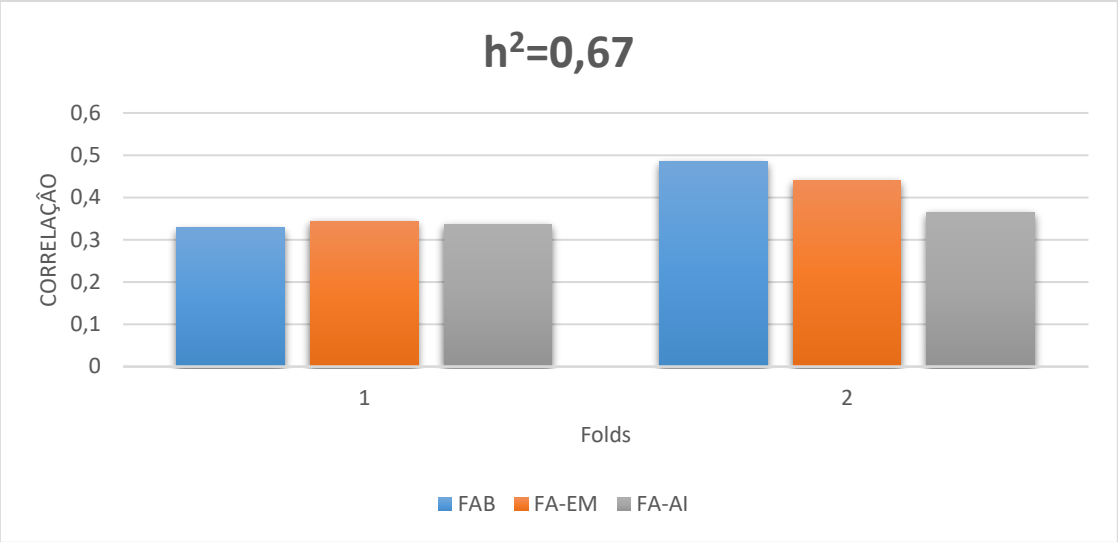


Figura 5-Gráfico de barras da correlação proveniente da validação cruzada considerando desbalanceamento de 50% de perda de genótipo nos ambientes usando o modelo FA bayesiano (FAB) FA padrão (FA-EM e FA-AI).

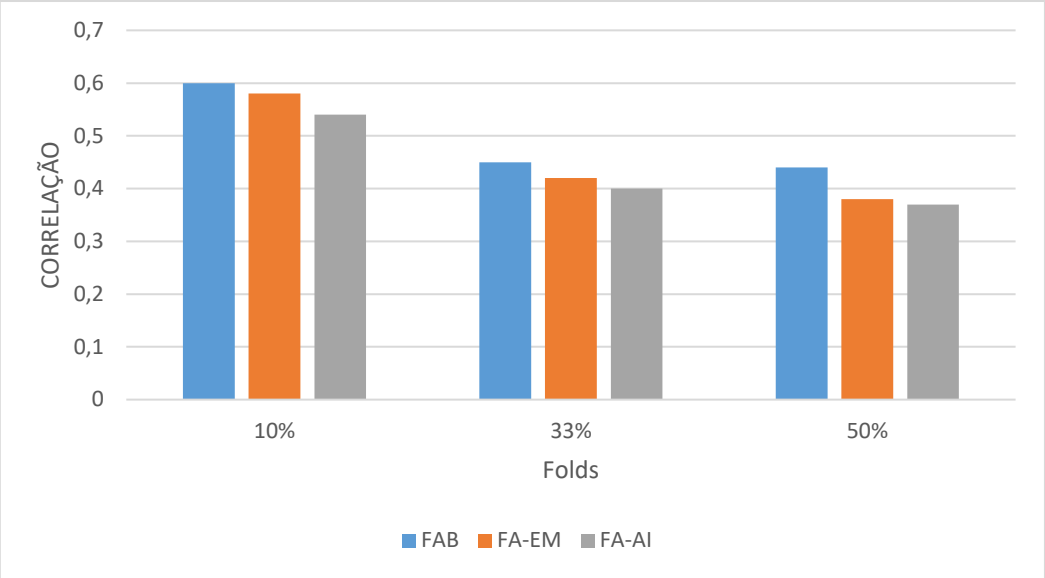


Na maioria dos folds o modelo FAB superou o FA-AI e FA-EM. Como pode se observar, a 10% o FAB teve baixa capacidade preditiva em relação ao modelo FA-EM nos folds 6 e 8, e praticamente igual a desses modelos fold 1 a 33% e 50% de perda de genótipos.

Quanto ao desbalanceamento a 33%, e a 50% praticamente todos os híbridos sofreram perdas em algum ambiente. Independentemente do nível de desbalanceamento aplicado, a grandeza dos valores de correlação ficou acima de 0,30 (Fig. 6), e modelo bayesiano foi o que teve as predições mais elevadas nos níveis de desbalanceamento considerado. Também é importante destacar que em todos os procedimentos os genótipos foram faltantes pelo menos uma vez em cada local.

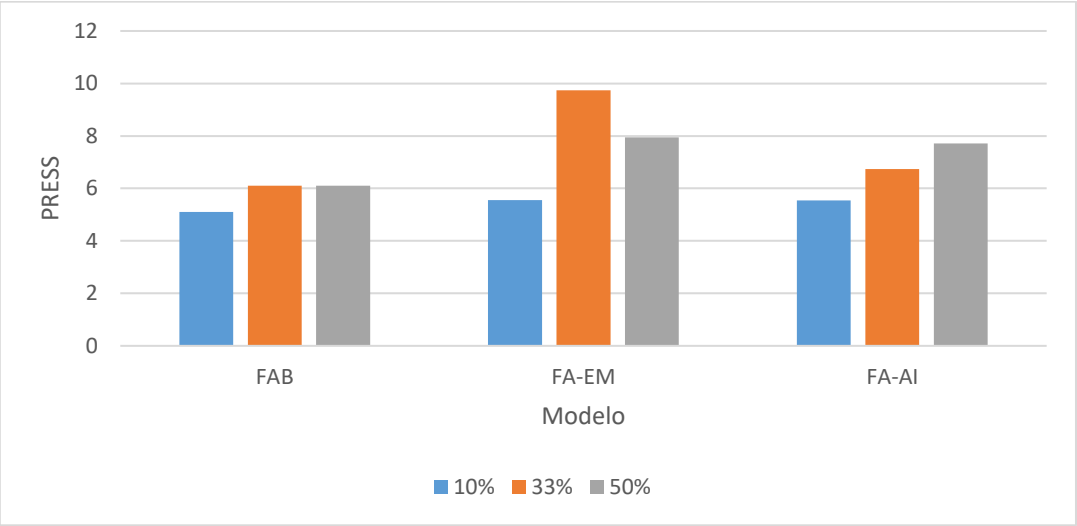
A capacidade preditiva do modelo para níveis de perdas de 33% e 50% de genótipos nos ambientes pode ser considerada praticamente à mesma nos modelos FA-EM e FA-AI. Porém, de forma geral, esses resultados indicam que mesmo em níveis de desbalanceamento entre 33% e 50% de perda de genótipos nos ambientes, o modelo FA continua sendo eficiente em predizer os valores fenotípicos (ou genotípicos) de híbridos perdidos não testados aos diversos ambientes com elevada acurácia.

Figura 6-Gráfico de barras da correlação proveniente da validação cruzada considerando os níveis de desbalanceamento de 10, 33 e 50% de perda de genótipos nos ambientes usando o modelo FA bayesiano (FAB) FA padrão (FA-EM e FA-AI).



A PRESS (Figura 7) a 10% foi a mais baixa comparada a dos outros níveis de desbalanceamento para os três modelos, embora a dispersão tenha sido praticamente a mesma. Em todos níveis de perda considerados o modelo FAB teve a menor PRESS, portanto o melhor modelo. A 10% os modelos FA-AI e FA-EM tiveram a mesma PRESS, portanto a escolha do modelo deve ser feita se baseando em outros critérios como custo computacional, mas a 33% o FA-AI seria o preterido que FA-EM e preferido a 50%.

Figura 7- Gráfico de barras da PRESS proveniente da validação cruzada considerando os níveis de desbalanceamento de 10, 33 e 50% de perda de genótipos nos ambientes usando o modelo FA bayesiano (FAB) FA padrão (FA-EM e FA-AI).



4.1.3. Seleção do modelo para predição

Com base neste estudo de simulação foi considerado um desbalanceamento de 10% e foram ajustados modelos FA5, FA4, FA3, FA2 e FA1 com cinco, quatro, três, dois e uma carga fatorial, respectivamente, para avaliar a capacidade preditiva de cada modelo com objetivo de selecionar o melhor. Como já referenciado, a seleção do modelo foi baseada no critério PRESS e eficiência estatística (SE) definida como a razão do modelo FA5 versus os demais.

As figuras 8 a 9 mostram os gráficos dos 5 modelos utilizados no estudo versus a correlação entre o valor fenotípico observado (simulado) e o fenotípico predito e a PRESS (alguns dos critérios usados para seleção de modelos). Observando a figura 8 verificamos que o modelo FA2 (com k=2) apresentou maior valor em relação aos demais indicando maior capacidade preditiva. Pelo critério PRESS o melhor modelo é o FA1, mas este modelo teve capacidade preditiva baixa quando comparado ao modelo FA2.

Figura 8- Ockham's plot referente ao modelo FAB versus correlação entre o valor fenotípico observado e predito (maior melhor).

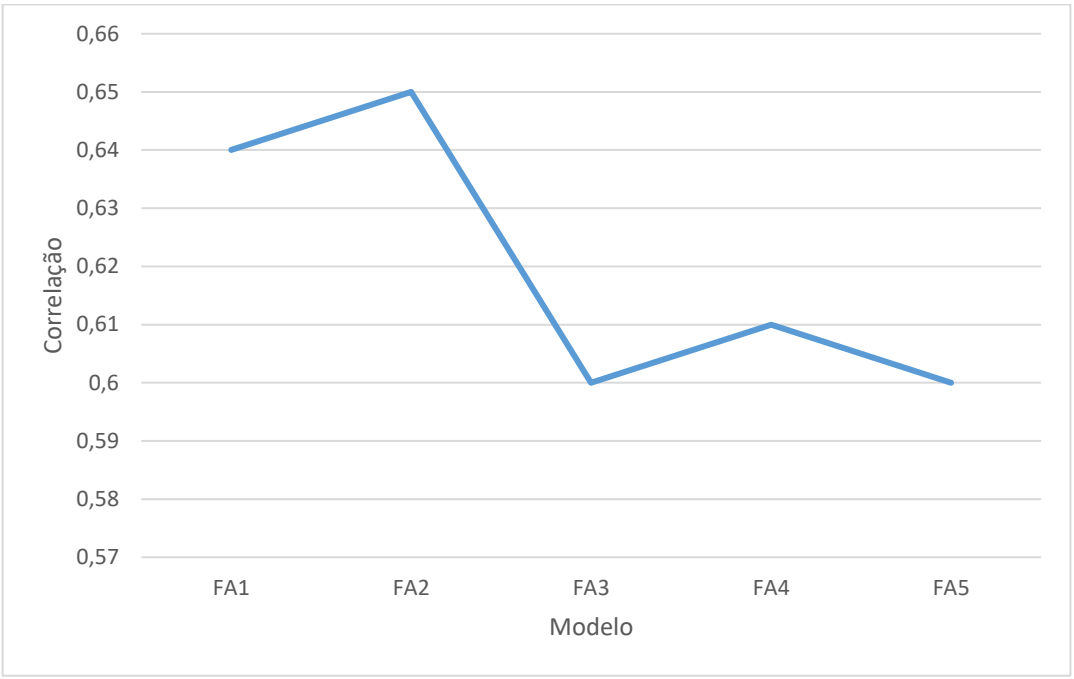
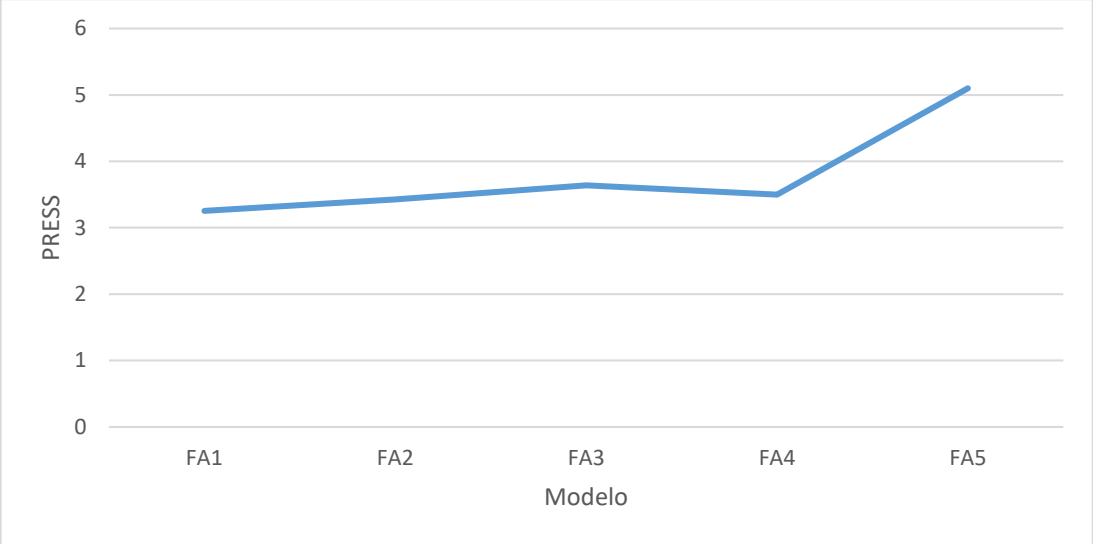
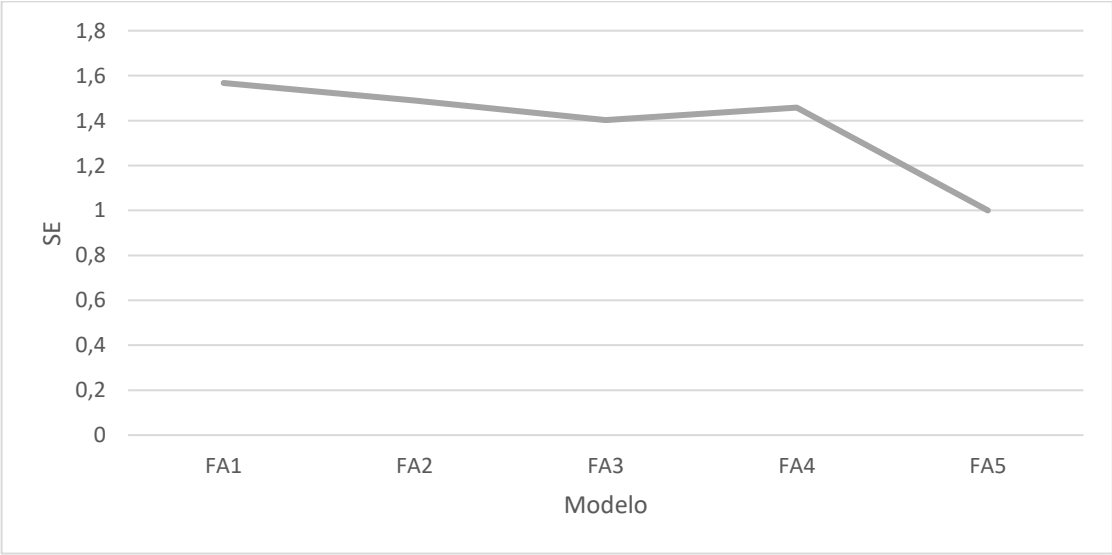


Figura 9-Ockham's plot referente ao modelo FAB versus prediction sum square- PRESS (menor melhor).



Na Figura 10 é apresentado o gráfico da eficiência estatística dos modelos em competição também. Nesse gráfico é notável que a eficiência diminua do FA1 ao FA3 onde encontramos um pico, demonstrando que uso de um k=1 seria suficiente para a representação dos dados. Porém, esse modelo teve um menor desempenho que FA2 quando o critério considerado foi o da correlação. Além disso, o FA2 geralmente é considerado o mais interpretável para os melhoristas de plantas, pois o primeiro eixo pode ser visto como representando a produção (adaptabilidade) e o segundo a estabilidade como no modelo SREG2/GGE (BURGUEÑO et al., 2008; STEFANOVA; BUIRCHELL, 2010).

Figura 10- Ockham's plot referente ao modelo FAB versus eficiência estatística-SE (maior melhor).



Ainda é possível verificar que o modelo FA4 apresentou maior eficiência que FA3, mas o modelo FA4 apresenta mais parâmetros que o modelo FA3, ou seja, em termos de parcimônia o modelo FA3 é preferido, mas considerando a eficiência estatística a PRESS, a correlação o FA4 seria o preferido.

4.2. Dados experimentais

Para ilustrar o desempenho do modelo FAB foi utilizado um conjunto de dados reais referente a um experimento em blocos incompletos, com 50 genótipos avaliados em 10 ambientes de testes, em presença de desbalanceamento. Foram simuladas 85 mil cadeias de Markov e, de modo análogo a (análise) anterior, descartou-se as primeiras 400 observações e a amostragem foi conduzida realizando-se saltos a cada 9 observações. Com a realização desses procedimentos foram selecionadas 9400 observações, para cada parâmetro. As convergências das cadeias foram verificadas pelos critérios de Raftery e Lewis (1992) teste de Heidelberger e Welch e teste de Gweke (GELMAN et al., 2013) e ainda pela observação e interpretação de gráficos que indicaram boas propriedades de convergência para todos os parâmetros do modelo.

As regiões HPD a 95% de probabilidade dos escores ambientais (cargas fatoriais) são mostradas na Tabela 3 e nos biplots nas figuras 11 e 12, para os genótipos e ambientes, respectivamente. A Tabela 3 mostra também as regiões HPD das variâncias residuais. Tanto a cargas fatoriais bem como as variâncias residuais, isto é, as médias a posteriori foram estimadas dentro dos intervalos de credibilidade. Verificando os ambientes nota-se que a amplitude das variâncias é baixa, o que poderia sinalizar a escolha de um modelo que desconsidere a heterogeneidade de variâncias. Como discutido na literatura os modelos mais realísticos são os que levam em consideração a heterogeneidade de variâncias entre ambientes.

Dos ambientes em avaliação o E8 foi o que apresentou maior variância residual e E2 a menor, o que pode ser observado na figura 12 em que a região elíptica do E8 é mais dispersa e a do E2 concentrada. De certa forma pode-se afirmar que a variância residual influencia a forma da elipse dos ambientes (região de credibilidade). Os que têm variâncias residuais baixas apresentam regiões mais concentradas e variância residuais altas apresentam elipses com maior amplitude. Como se sabe o ângulo da elipse é determinado pela covariância residual. Neste caso, a covariância é zero, de modo que os dados não estão correlacionados, resultando em uma elipse alinhada com eixos. Além disso, é claro que as grandezas dos eixos de elipse dependem da variância. No nosso caso, de maior variância é na direção do eixo X, enquanto a menor variância está na direção do eixo Y.

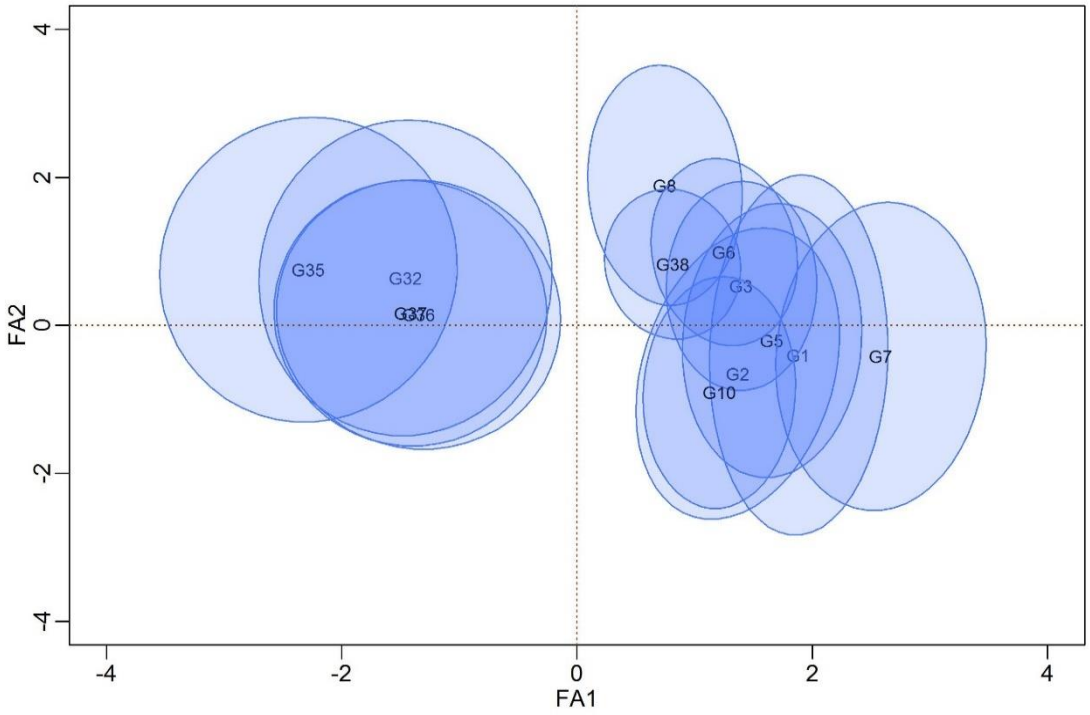
Tabela 4-Médias a posteriori (MP), regiões de credibilidade (I.C. 95%, LI: limite inferior, LS: limite superior) para as duas primeiras cargas fatoriais ambientais ($\tau_{j1}-\tau_{j2}$) e variância residual ($\sigma^2_{e_1}-\sigma^2_{e_{10}}$).

			HPD 95%					HPD 95%	
Par.	MP	sd	LI	LS	Par.	MP	sd	LI	LS
τ_{11}	0,190	0,178	-0,179	0,516	τ_{12}	0,070	0,231	-0,432	0,477
τ_{21}	0,625	0,179	0,283	0,980	τ_{22}	0,533	0,188	0,106	0,887
τ_{31}	0,291	0,151	0,002	0,591	τ_{32}	0,197	0,207	-0,242	0,568
τ_{41}	0,091	0,140	-0,195	0,359	τ_{42}	0,340	0,169	-0,015	0,664
τ_{51}	0,600	0,176	0,262	0,953	τ_{52}	0,253	0,182	-0,118	0,606
τ_{61}	0,794	0,199	0,379	1,163	τ_{62}	-0,216	0,310	-0,761	0,454
τ_{71}	0,944	0,200	0,547	1,327	τ_{72}	-0,386	0,216	-0,810	0,066
τ_{81}	1,223	0,234	0,760	1,665	τ_{82}	0,136	0,277	-0,426	0,657
τ_{91}	0,440	0,150	0,134	0,720	τ_{92}	-0,184	0,179	-0,552	0,165
τ_{101}	0,693	0,195	0,311	1,074	τ_{102}	-0,227	0,261	-0,707	0,328
$\sigma^2_{e_1}$	3,4642	0,2869	2,9392	3,996					
$\sigma^2_{e_2}$	1,4465	0,2187	1,2076	1,6734					
$\sigma^2_{e_3}$	1,9769	0,1655	1,6827	2,2918					
$\sigma^2_{e_4}$	1,6672	0,1634	1,3776	1,9581					
$\sigma^2_{e_5}$	3,2711	0,303	2,7526	3,7986					
$\sigma^2_{e_6}$	1,8623	0,1538	1,585	2,1727					
$\sigma^2_{e_7}$	2,1392	0,1968	1,8324	2,4674					
$\sigma^2_{e_8}$	4,0689	0,2984	3,4846	4,6402					
$\sigma^2_{e_9}$	1,6871	0,1472	1,409	1,9773					
$\sigma^2_{e_{10}}$	3,3718	0,2624	2,8713	3,8595					

Par=parâmetro

A Figura 11 mostra o gráfico dos escores genotípicos a posteriori, onde as áreas interiores sombreadas do gráfico são regiões posteriores bivariadas HPD a 95% de probabilidade. Apenas genótipos que não incluem a origem foram representados. Do mesmo modo, a figura 13 mostra o biplot os escores ambientais para ambientes (E2, E4, E5, E6, E7, E8, E9 e E10) que não incluíam a origem (0,0) em suas regiões HPD a 95% de probabilidade.

Figura 11 -Regiões de credibilidade a 95% de probabilidade para os escores fatoriais genotípicos que não englobam a origem, usando abordagem bayesiana.



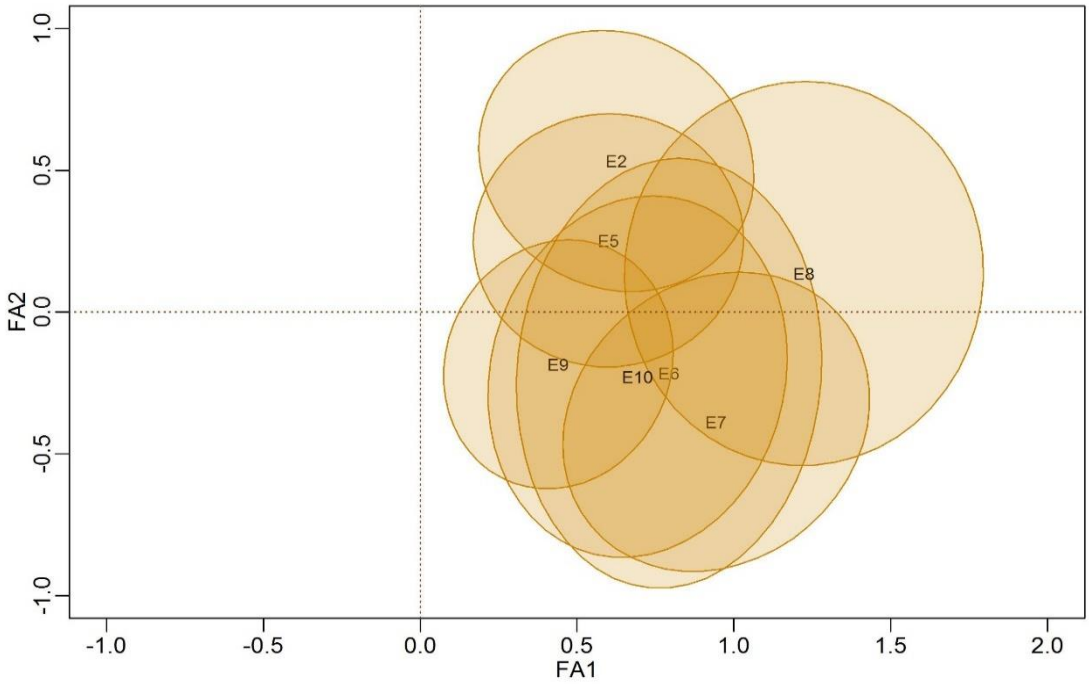
Verificou-se uma correlação entre o escores genotípicos e a produção foi de (r^2=0,897) no modelo bayesiano e (r^2=0,929), que justificaria uma interpretação dos biplot similar aos modelos GGE-biplot.

Uma vez que existe uma clara sobreposição entre os escores fatoriais dos genótipos G32, G35, G36 e G 37 nas regiões HPD (Figura 11), a 95% de probabilidade, e todas as regiões são cortadas pelo primeiro eixo, estes genótipos não produziram diferenças estatísticas nos efeitos GEI e na produtividade. Os genótipos localizados à direita do eixo FA2, por outro lado, possuem produtividade maior que a média geral e nesse requisito seriam os mais interessantes. Observamos, por exemplo, que o subgrupo {G1, G2, G5, G6, G10, G7} é formado por genótipos que possuem resultados semelhantes, tanto com relação à produtividade como com relação ao efeito da interação, pois as regiões se interceptam (produtividade semelhante) e todas são cortadas pelo primeiro eixo no biplot (genótipos estáveis).

Os escores ambientais para E2, E4, E5, E6, E7, E8, E9 e E10 (Tabela 3; Fig. 12) não incluem a origem (0,0) nas regiões HPD, a 95% de probabilidade, indicando que os efeitos GEI induzidas nesses ambientes foram significativos. Nesse biplot observam-se diferentes níveis de sobreposições entre as regiões de credibilidade, evidenciando a dificuldade de separar subgrupos distintos com relação ao efeito da interação. Diferença mais expressiva pode ser observada apenas entre E2 e E7.

A resposta conjunta de genótipos e ambientes pode ser examinada ao se considerar as Figuras 11 e 12 simultaneamente. Os genótipos, G32, G35, G36 e G37 formam um subgrupo, enquanto os genótipos G1, G2, G3, G5, G6, G7, G8, G10 e G38 formaram outro subgrupo distinto, no que se refere à produtividade. Naturalmente {G32, G35, G36, G37} não seriam interessantes para recomendação.

Figura 12-Regiões de credibilidade a 95% de probabilidade para as cargas fatoriais de ambientes cujas regiões não englobam a origem, usando abordagem bayesiana.



As Figuras 13 e 14 apresentam uma breve comparação dos resultados de biplots obtidos a partir do ajuste do FA pelo método FA-AI com duas cargas e o modelo Bayesiano proposto. Ambos os biplots mostram os genótipos, {G32, G35, G36, G37} e genótipos {G1, G2, G3, G5, G6, G7, G8, G10, G38} em diferentes quadrantes opostos, como os que contribuem mais para o GEI. Nas Figuras 13 e 14, verifica-se que as estimativas dos dois modelos foram coiscidentes, tanto para os escores genotípicos bem como para cargas. Semelhanças podem ser examinadass quando se compara a distribuição dos esscores ambientais nas análises do biplot padrão (Fig. 14), em comparação com aqueles obtidos a partir do modelo bayesiano com incerteza (Fig. 12). É possível verificar a alteração da posição do E8 do modelo FA-AI (quarto quadrante) para o FA bayesiano (segundo quadrante), ademais no modelo FA bayesiano é muita clara a identificação de ambientes que causam a interação se considerar a incerteza no biplot.

Figura 13- Análise biplot dos escores genotípicos usando modelo FA-AI (Vermelho) FAB (Azul), considerando 50 genótipos avaliados em 10 ambientes.

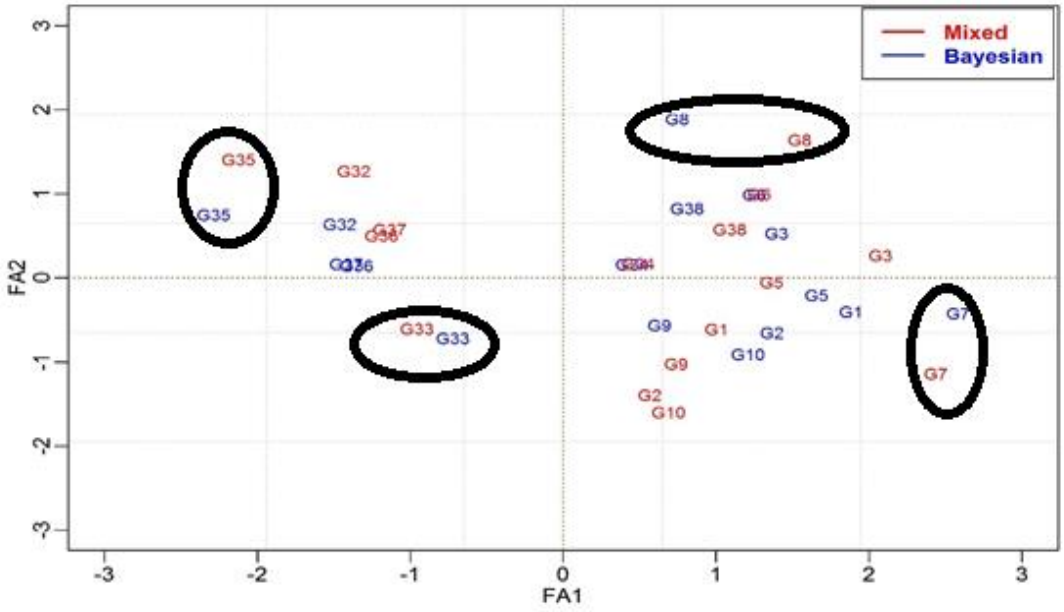
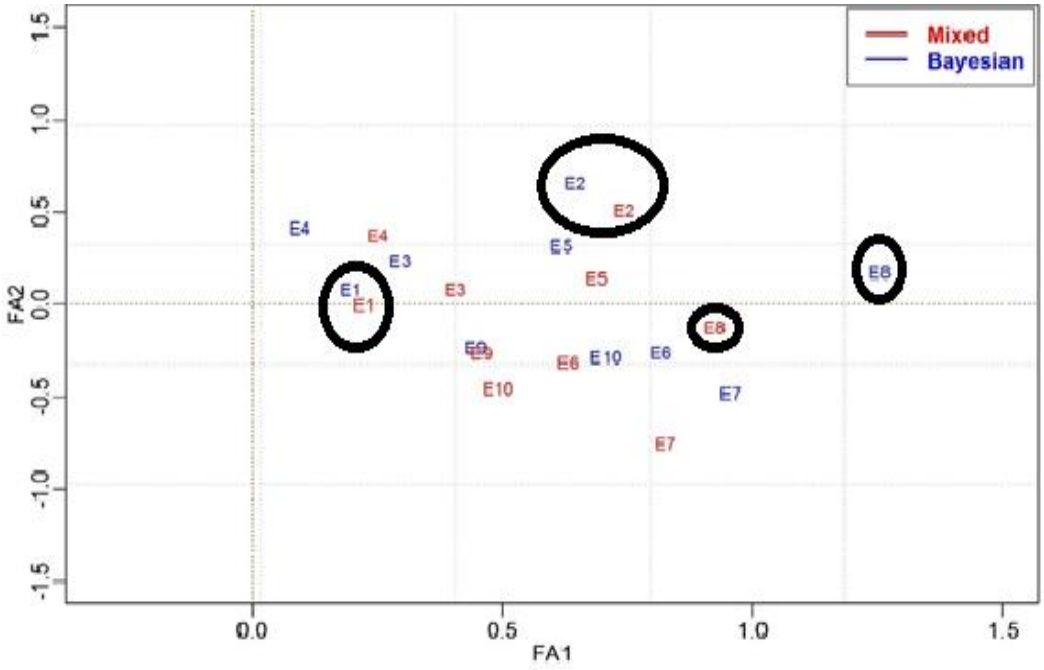


Figura 14-Análise biplot dos escores ambientais usando modelo FA-AI (Vermelho) e FAB (Azul), considerando 50 genótipos avaliados em 10 ambientes.



4.2.1. Seleção do modelo

Na análise de fatores bayesiana uma das questões crucias a ser respondida além da escolha das distribuições a priori para hiper-parâmetros, é a escolha do número adequado de fatores k a serem retidos no modelo. Neste trabalho foi proposto o uso do critério Akaike Monte

Carlo (AICM) e a $\Delta AICM$ -diferença entre o valor do AICM dos modelos em competição versus o modelo completo, propostos por Raftery et al. (2007) para a seleção do número de fatores k.

Na Tabela 5 apresentamos os valores da AICM da log verossimilhança posteriori dos 10 modelos em competição usando dados reais, verificando-se que pelo critério AICM o melhor modelo é o completo -FA10 seguido do FA2.

Tabela 5- Valores de AICM e a ΔAICM-diferença entre o AICM do modelo completo e os demais, respectivamente para seleção de modelos FAk, (k=1...,10) e o ranqueamento dos modelos.

Modelo	AICM	Rank	ΔAICM	Rank
FA1	-4,351	1º-FA10	-0,007	1º -FA2
FA2	-4,347	2º -FA2	-0,003	2º FA4
FA3	-4,351	3º- FA4	-0,007	3º -FA1
FA4	-4,35	4º -FA1	-0,006	4º -FA3
FA5	-4,351	5º -FA3	-0,007	5º -FA5
FA6	-4,351	6º -FA5	-0,007	6º -FA6
FA7	-4,354	7º -FA6	-0,01	7º -FA9
FA8	-4,354	8º -FA9	-0,01	8º -FA7
FA9	-4,353	9º- FA7	-0,009	9º -FA8
FA10	-4,344	10º -FA8	-	-

Muitas vezes estamos interessados em selecionar um modelo mais parcimonioso, comparando os modelos em competição contra o completo. Por essa tarefa não ser trivial quanto parece em busca de alternativas para seleção de modelo nesse trabalho usamos o critério ΔAICM. Na tabela 4 é possível verificar ainda a mudança em relação ao modelo completo, com valores menores indicando o modelo preferido. Por esse critério, verificamos que o modelo FA2 é o melhor em relação aos demais seguido do FA4 e pelo FA1. Em comparação com a simulação é notável a coincidência na seleção do melhor modelo.

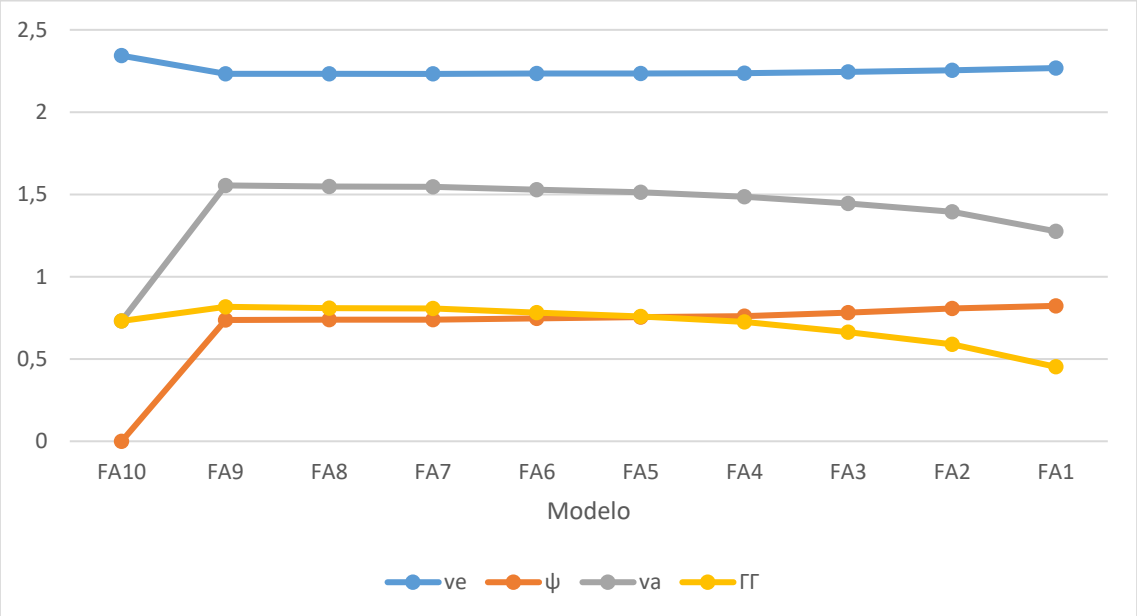
Como foi verificado na Tabela 4, é difícil chegar ao melhor modelo pelos critérios de informação, pois a diferenciação só ocorre na terceira casa decimal, tornando- se necessário uma medida adicional para seleção do modelo. Com eeste trabalho foi observado o comportamento dos componentes de variância o que está apresentado na, Figura 15.

Nos modelos FA, a covariância genética é estimada como ($\boldsymbol{\Sigma} = \left(\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}\right)^{k-1} + \boldsymbol{\Psi}$), e como forma de garantir a identificabilidade em modelos FAk-1 a matriz de cargas é estimada como ($\boldsymbol{\Sigma} - \boldsymbol{\Psi}$).

Na figura 15 pode-se verificar que quando se ajusta o modelo completo-FAk a proporção da variância genética (va - $\sigma_g^2 = \sqrt[10]{\left(diag\left|\boldsymbol{\Gamma}\boldsymbol{\Gamma}^{\mathrm{T}}\right|\right)}$ é totalmente recuperada pelas cargas

($\sqrt[10]{diag\left(\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\right)}$) e com uma variância residual (ve $\sigma_e^2 = \sqrt[10]{|\mathbf{R}|}$) maior. Quando se retira um eixo do modelo completo podemos verificar a diminuição da variância residual e aumento da variância genética, que pode ser explicado pelo fato de ao reduzir os eixos, o número de parâmetros do modelo decai, e parte da variância residual ser transferida a variância específica do modelo, que passa a ser recuperada pelas cargas e variância específicas ($\left(\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\right)^{k-1} + \boldsymbol{\Psi}$). A de destacar que, como parte da variância residual passa a ser explicada pela variância, específica carece de estudos. Pode-se verificar que do modelo FAk-1 ao modelo FA1 as variâncias σ_e^2 e ψ ($\sqrt[10]{\boldsymbol{\Psi}}$) quase se mantém constantes, acompanhados do decaimento da $\sqrt[10]{diag\left(\boldsymbol{\Gamma\Gamma}^{\mathrm{T}}\right)}$ e aumento de ψ. É possível notar que σ_g^2 a medida que vão se retirando os eixos passa ser estimada como $\sigma_g^2 = \sqrt[10]{\left(diag\left|\boldsymbol{\Gamma\Gamma}^{\mathrm{T}} + \boldsymbol{\Psi}\right|\right)}$. Nota-se que a mudança na ψ de FA1 para FA2 é quase nula, como acontece na σ_e^2. Embora σ_g^2 de FA1 para FA2 tenha um ligeiro aumento e estejam acima de σ_g^2 do FA10, pode-se verificar que o uso de modelos complexos (com mais eixos) não traz nenhuma vantagem ou ganho no cálculo da σ_g^2 ou σ_e^2. Dessa forma a realização de testes de seleção de modelos pode ser desnecessária.

Figura 15-Comportamento da Variância residual (ve), genética (va) e específica (ψ) nos modelos diferentes modelos.



4.3. DISCUSSÃO

Modelos FA vêm sendo usados com frequência em programas de melhoramento para a análise da interação GEI (BURGUEÑO et al., 2008; CROSSA, 2012; KELLY et al., 2007; MEYER, 2009), em estudos de seleção genômica (OAKEY et al., 2016), em análise sensórios (SMITH et al., 2003) e em psicologia (KARATZA, 2006). Esses modelos oferecem diversas vantagens quando comparados a métodos tradicionais, sobretudo sobre aqueles que consideram os parâmetros como sendo de efeitos fixos. Contudo, um grande problema relacionado à abordagem FA tem sido a falta de identificabilidade, na estimação dos parâmetros.

O objetivo desta pesquisa foi propor uma abordagem bayesiana ao modelo FA útil a pesquisadores de forma geral. Podem-se mencionar algumas vantagens das estimativas condicionais a posteriori do modelo FAB sobre o FA-misto. Por exemplo, o fato de todas as quantidades desconhecidas serem aleatórias na ótica bayesiana evita o problema de determinar quais os efeitos devem ser considerados fixos e/ou aleatório. Além disso, fornece um método natural para derivar regiões de confiança em torno dos parâmetros genotípicos e ambientais e, sobretudo, sobre aqueles que descrevem da GEI, permitindo identificar grupos de genótipos e ambientes semelhantes, como assinalado por Yang et al. (2009). Uma desvantagem importante do FA-misto e que este não oferece nenhuma inferência estatística para identificar grupos separáveis.

A principal diferença do modelo aqui proposto para a abordagem clássica reside, primeiramente, em um menor número total de fatores latentes estimados ($k<<p$). O modelo proposto estima os parâmetros de $\boldsymbol{\Sigma}$ de forma similar de um modelo com variância não estruturada (UN), ou, seja $p(p+1)/2$ parâmetros. Outra vantagem seria na escolha do k. A escolha de k é um problema muito difícil, não resolvido, e escolhas inadequadas podem resultar em estimativas tendenciosas e instáveis de $\boldsymbol{\Sigma}$ e $\mathbf{R}$ (MEYER; KIRKPATRICK, 2008; RUNCIE; MUKHERJEE, 2013). Em nosso modelo, permitimos incorporação de todos os fatores latentes, ou seja, a estimação de todas as cargas, apesar de assumirmos que a maioria deles é relativamente sem importância. Esta sutil diferença é importante porque remove a necessidade de escolher com precisão k, enfatizando a estimativa da magnitude de cada fator latente.

Em relação a abordagens envolvendo inferência bayesiana, nosso método se diferencia pela construção do modelo, que é baseada em decomposição espectral da matriz de covariâncias genética para garantir identificabilidade do modelo, evitar a ocorrência de casos Heywood e a necessidade de rotacionalidade das cargas. Neste estudo assumimos independência entre os

ambientes e, assim, para cada componente de variância residual atribuiu-se priori qui-quadrada escalada invertida. Caso não se queira assumir tal independência, pode-se atribuir uma priori inversa Whishart para a matriz de variância residual, como em Campos e Gianola (2007). Para o modelo proposto por Runcie e Mukherjee (2013) que foi baseado no modelo de Campos e Gianola (2007) a esparcidade assumida nas matrizes de cargas como priori.

Deve-se destacar que, como forma de garantir a indentificabilidade do modelo de Campos e Gianola (2007) aplica uma vetorização a matriz de cargas, e Runcie e Mukherjee (2013) induzem a esparcidade a matriz de cargas, além de colocarem uma hierarquia a cada elemento da matriz de cargas, ou seja, uma priori hieraqruica. Vale ressaltar ainda que a questão de rotacionalidade de cargas não é discutida no trabalho de Campos e Gianola (2007) e Runcie e Mukherjee (2013) afirmam que a rotacionalidade é garantida por meio da imposição de restrições na matriz de cargas que é formalizada na especificação da priori (distribuição hierárquica). Como se sabe a rotacionalidade da matriz de cargas não influencia a estimação dos parâmetros do modelo, mas pode dificultar as intrepretações biológicas (SMITH; CULLIS; THOMPSON, 2001) e dificultar a avaliação da convergência MCMC (WEST, 2003).

A principal vantagem do modelo FAB aqui proposto sobre os demais FAB reside no fato do modelo fornecer estimativas robustas de parâmetros de covariância, das cargas e dos escores fatoriais, sem necessidade de rotacionalidade das cargas, além da sua capacidade de produzir estimativas interpretáveis sem ocorrência de casos Heywood, como observados em Smith, Cullis e Thompson (2001). A principal limitação desse método é o tempo computacional, que aumenta com o número de caraterísticas/ambientes analisados.

Na estimação dos parâmetros constatou-se que as cadeias MCMC geradas para as coordenadas dos vetores singulares, a partir do segundo, convergiram para duas soluções, iguais em módulo e, consequentemente os escores fatoriais (genotípicos) estão sujeitos a mudança de sinal, devido à combinação linear existente entre as cargas e os escores fatoriais de Oliveira et al. (2015) relata que esse problema de mudança de sinais observado é arbitrário, de modo que as cadeias devem ser separadas para as duas soluções e as subcadeias com mesma solução agrupadas para a realização do teste de convergência e obtenção da amostra para o processo de inferência sobre os parâmetros. A não observação desse fato pode levar a estimativas equivocadas, fazendo com que os valores estimados das coordenadas, a partir do segundo vetor, se aproximassem do valor zero, o que obviamente compromete os resultados e interpretações da adaptabilidade e estabilidade.

Neste estudo verificamos um alto desempenho do modelo em cinco ambientes simulados, bem como com os dados reais em 10 ambientes (os resultados se mostraram

robustos). Constatou-se a eficácia da abordagem visto que o padrão da representação biplot, da análise convencional, foi conservado pelo biplot composto pelas médias a posteriori. Além disso, como observado nas Tabelas 1, 2 e 3, os valores obtidos para as médias a posteriori dos efeitos ambientais, variância residual são próximos dos verdadeiros e as coordenadas dos escores dos dois primeiros eixos são bem próximos às estimativas de máxima verossimilhança restrita (REML) do modelo FA-AI, o que também indica que as prioris assumidas na nossa abordagem se aproximam da abordagem do modelo FA-AI, mas sem necessidade de imposição de rotacionalidade como já referenciado.

Outras propostas para análise de dados multiambientes podem ser encontradas na literatura, dentre as quais se destacam o uso dos modelos AMMI ou SREG bayesianos (CROSSA et al., 2011; JARQUÍM et al., 2016; OLIVEIRA et al., 2015, 2016; PEREZ-ELIZALD; JARQUIN; CROSSA, 2011). Uma limitação presente nessas abordagens, em relação a FAB, está em assumir homogeneidade de variâncias como pressuposto, bem como as demais pressuposições de modelos baseados em análise de variância, embora se saiba que esses modelos são flexíveis para lidarem com esses problemas. Ao assumirem homogeneidade de variâncias, a amostragem dos vetores singulares é realizada em hiperesferas por meio de distribuições Von Mises-Fisher, como proposto por Liu (2001). Aqui, como assumimos variâncias diferentes não é possível obter Von Mises-Fishers como condicionais a posteriores e sim normais multivariadas, de forma que a amostragem dos vetores é realizada em hiperelipses e os mesmos são colocados no subespaço correto por transformação ortogonal, atendendo assim as restrições impostas pela decomposição espectral.

Embora o objetivo com esse estudo não fosse comparar a capacidade preditiva do FAB versus FA misto, os resultados mostram que usando diferentes tipos de FA mistos (AI e EM) o FAB se mostrou melhor em todos os níveis de perdas de genótipos considerados. A perda de previsibilidade, isto é, correlações decrescentes do FAB para o FA-EM e FA-AI foi cerca de 3% e 5% com desbalanceamento de 33% de ¨6% e 7% com desbalanceamento de 50%, respectivamente. Utilizando o h^2 médio esperaríamos uma correlação entre o valor observado e o predito de 0,82, mas considerando os k-folds cada nível de perda, por exemplo, a 10% verificou-se uma correlação de 0,86 no modelo bayesiano, uma evidência que o limite teórico poderia ser superado por esse modelo. Em alguns k-folds podemos verificar a perda marginal do FAB em relação ao FA-EM exemplo a 10% e capacidade preditiva igual aos FA-EM e FA-AI aos 33% e 50%. O ganho relativo do FA-EM em relação ao FA-AI pode ser explicado pela estimação dos modelos, segundo Nuvunga et al. (2015) o FA-EM é estimado em dois estágios, estimando no primeiro estágio a variância UN que é completa e aproximado no segundo estágio

ao FA e requer estimação de todo parâmetros do modelo o que pode levar a perda de informação, enquanto que o FA-AI estima a matriz de variância de forma aproximada e com menos parâmetros em único estágio. De forma geral, o estudo mostra que quando se usa o modelo FAB, mesmo com elevados níveis de perda, a sua predição é superior ao FA misto e a previsibilidade (correlações) é substancialmente aumentada. Assim, os resultados deste estudo mostram que a modelagem GEI por FAB pode ser muito melhor que o modelo FA mistos.

Como discutido em Crossa (2012) a montagem de modelos FA pode se dar semelhante aos modelos SREG (se o efeito de G for confundido com GEI) ou ao modelo AMMI (se o efeito de G for excluído da GEI) o que, de acordo com Burgueño et al. (2012), não tem influência clara em termos de predição. Contudo, deve-se salientar que os modelos FA (SREG), conforme apresentado aqui, são mais parcimoniosos do que os modelos que ajustam o efeito de genótipos separadamente da GEI e, portanto, mais fácil de ajustar. O conjunto de dados, usado neste estudo, é relativamente pequeno, em comparação a outros com maior número de genótipos, tais como, aqueles usados em ensaios iniciais de melhoramento. Para ilustrar a previsibilidade dos vários modelos estatísticos, em diferentes graus de complexidade de GEI e para avaliar a robustez do modelo FA foi feita a análise de experimentos com diferentes níveis de desbalanceamento usando estrutura GEI complexas (BURGUEÑO et al., 2012; KELLY et al., 2007; NUVUNGA et al., 2015). Contudo, o nosso conjunto de dados parece ter sido suficiente para os principais propósitos do estudo.

Na seleção de modelos, em dados simulados, foram adoptados três critérios como complementares: o critério PRESS, a correlação entre o valor fenotípico observado e o predito e a eficiência estatística (SE). A opção pelo uso dos três critérios deveu-se ao fato de diferentes critérios tenderem a escolher modelos diferentes, como complementares espera-se que um compense a limitação do outro. É notável que o modelo escolhido pela PRESS e SE foi o mesmo, mas tomando em conta a correlação, o modelo foi diferente, cabendo à sensibilidade do pesquisador na escolha do modelo a ser adotado. A PRESS é um critério útil para seleção de modelos quando se tem dados balanceados. Nos dados experimentais não seria a melhor opção, já que os dados experimentais são naturalmente desbalanceados de certa maneira. Ademais, vários critérios de teste para determinar o melhor modelo encontram-se disponíveis na literatura. Estudos de simulação, examinando a sua utilidade, no entanto, geralmente têm rendido resultados não muito consistentes, tanto entre diferentes testes, como na capacidade de encontrar o melhor modelo ou critério (MEYER, 2009; MEYER; KIRKPATRICK, 2008). Em modelos mistos, a seleção de modelos com base na log verossimilhança, critérios de informação ou fatores de Bayes são uma escolha óbvia. Mas em modelos bayesianos que envolvem MCMC

esta prática tem se tornado pouco atraente. Por exemplo, o cálculo direto do fator de bayes usando amostra MCMC requer uma demanda computacional maior, o que motivou a nossa escolha pelo critério AICM proposto por Raftery et al. (2007) e o desempenho dos componentes de variância.

Na nossa abordagem, o uso de diferentes critérios de informação tende a escolher diferentes modelos (Tabela 4) e o AICM tende a escolher o modelo completo que em tese explicaria toda variância genética e o ΔAICM escolhe o modelo FA2, que é o mais parcimonioso e com justificativa genética (BURGUEÑO et al., 2007; STEFANOVA; BURCIIEL, 2010; NUVUNGA et al., 2015), além de maior capacidade preditiva, como verificado por Burgueño et al. (2008) e Kelly et al. (2007). O uso de critérios de informação como método de seleção de número de fatores (seleção de modelo em análise fatorial) possui várias ambiguidades e é sempre sujeito a críticas, pois diferentes critérios tendem a selecionar modelos diferentes, como pode ser observado na tabela 4 e em Smith et al. (2015). Pode-se verificar ainda na tabela 4 que a diferenciação dos modelos só é verificada na terceira casa decimal. Smith et al. (2015) destacam a necessidade de uso de medidas adicionais além de critérios de informação, na seleção de modelos como a proporção de variância explicada pelos componentes. Como vimos que em modelos FA é necessário que se garanta a indentificabilidade. A medida que o modelo k-1, é ajustado, a variância residual diminui e a genética aumenta, que passa a ser explicada pela cargas e variância específica (Figura 15). Do modelo k-2, k-3, ...,1 é notável que a variância específica e a residual se mantêm constantes, ilustrando que a aplicação de critérios de informação, como método de seleção de modelos, não é fácil de ser implementada.

A escolha do modelo FA2 está de acordo com vários resultados anteriores relatadas por outros pesquisadores. O aumento do número de componentes por encaixe, digamos, a modelos FA3, FA4 e FA5, não garante o aumento da capacidade de predição, mas certamente aumentará a complexidade do modelo, e é duvidoso que ele vai produzir melhor ajuste. Os resultados obtidos por Burgueño et al. (2007, 2008), Kelly et al. (2007) e Nuvunga et al. (2015) mostram que os modelos FA com mais de dois componentes melhoraram as estimativas de variância-covariância, mas isso não se reflete nos valores preditos de genótipos (EBLUPs). Como é possível verificar na Tabela 8, a correlação entre o valor predito e observado, derivados da simulação, foram maiores quando se usou o FA2, mas podemos verificar que há perda de predição para o modelo FA1 (de 1%), melhor modelo selecionado pelo critério PRESS e SE. Portanto, o modelo FA2 deveria ser o preferido, além de ser o mais facilmente interpretado (ou

seja, com interpretação desejável aos melhoristas como a feita para o modelo SREG2/GGE), como já enfatizado anteriormente.

Regiões elípticas de credibilidade (a 95% de probabilidade) foram representadas no biplot (Fig. 1, 2, 11 e 12). Essas regiões são baseadas nas distâncias euclidianas dos pontos em relação ao centro das distribuições dos escores, semelhantes às regiões apresentadas por Oliveira et al. (2015). Jarquím et al. (2016) e Oliveira et al. (2016) apresentaram regiões de credibilidade HPD para os escores genotípicos e ambientais, para modelos SREG bayesianos em que o efeito de genótipos é modelado conjuntamente com GEI e mostraram como separar grupos homogêneos de genótipos e ambientes com relação ao efeito da interação. Métodos para obter regiões de confiança livres de pressupostos teóricos com relação à distribuição dos escores também podem ser encontrados em Hu e Yang (2013b). É importante ressaltar que a inferência bayesiana oferece um método flexível e paramétrico para obtenção de inferência no biplot, baseado na distribuição conjunta a posteriori, ao contrário de métodos paramétricos frequentista difíceis de serem estendidos a modelos mais complexos e que requerem suposições restritivas sobre as distribuições de escores individuais (HU; YANG, 2013a; YANG et al., 2009). Aliás, no nosso modelo, foi verificado que a forma da elipse é fortemente influenciada pela variância residual, podendo se identificar ...por meio ... da elipse quais ambientes apresentam maior variabilidade.

Por fim é necessário salientar que o verdadeiro problema que os melhoristas enfrentam é predizer de forma independente o desempenho genotípico em outros locais (ou seja, locais não incluídos no MET) e anos futuros. Este estudo respondeu parcialmente a essa pergunta, mas não deu uma avaliação de predição independente quando todos os genótipos são completamente ausentes de todo o conjunto de locais.

5. CONCLUSÃO

Uma característica única do modelo proposto é que os fatores genéticos e ambientais são estimados em conjunto, em vez de separadamente, como em modelos de multiníveis de análise de fatores clássicos.

O modelo proposto produz estimativas robustas.

Os parâmetros do modelo foram estimados adequadamente, sendo identificável, sem necessidade de rotacionalidade das cargas fatoriais ou de imposição de restrições.

Em todos os modelos testados verificou-se que o modelo fator analítico com duas cargas (FA2) é o melhor na predição de valores faltantes e na análise de dados multi ambientes (MET).

Na análise de dados reais a seleção de modelos (ou escolha de número de fatores) não deve ser baseada somente em critérios de informação.

REFERÊNCIAS

AKAIKE, H. Factor analysis and AIC. **Psychometrika**, Williamsburg, v. 52, n. 3, p. 317-332, 1987.

AKAIKE, H. Information theory and an extension of the maximum likelihood principle. In: INTERNATIONAL SYMPOSIUM ON INFORMATION THEORY, 2., 1973, Budapest. **Proceedings…** Budapest: Akademiai Kiado, 1973. p. 267-281.

ANDERSON, T. W.; RUBIN, H. **Statistical analysis in factor analysis**. Stanford: Stanford University California Applied Mathematics and Statistics Labs, 1955. 84 p. (Report, TR30).

BARNETT, A. G. et al. Selecting the correct variance-covariance structure for longitudinal data in ecology: a comparison of Akaike, quasi-information and deviance information criteria. **Methods in Ecology and Evolution**, London, 2010. Disponível em: <http://eprints.qut.edu.au/19195/>. Acesso em: 5 ago. 2015.

BARTHOLOMEW, D. J.; KNOTT, M. **Latent variable models and factor analysis**. 2nd ed. New York: E. Arnold, 1999. 294 p. (Kendalls Library of Statistics, 7).

BAUER, A. et al. Bayesian prediction of breeding values by accounting for genotype-by-environment interaction in self-pollinating crops. **Genetics Research**, London, v. 91, p. 193-207, 2009.

BERGER, J. O.; GHOSH, J. K.; MUKHOPADHYAY, N. Approximation and consistency of Bayes factors as model dimension grows. **Journal of Statistical Planning and Inference**, Amsterdam, v. 112, p. 241-258, 2003.

BERGER, J. O.; PERICCHI, L. R. The intrinsic bayes factor for linear models. In: INTERNATIONAL MEETING ON BAYESIAN STATISTICS, 5., 1996, Oxford. **Proceedings…** Oxford: Oxford University Press, 1996. p. 25-44.

BERNARDO, J. M. Intrinsic credible regions: an objective Bayesian approach to interval estimation. **Test**, Madrid, v. 14, p. 317-384, 2005a.

BERNARDO, J. M. Reference analysis. In: DEY, D. K.; RAO, C. R. (Ed.). **Handbook of statistics 25**. Amsterdam: Elsevier, 2005b. p. 17-90.

BERNARDO, J. M.; SMITH, A. F. M. **Bayesian theory**. Chichester: Wiley, 1994. 610 p.

BUJA, A.; EYUBOGLU, N. Remarks on parallel analysis. **Multivariate Behavioral Research**, Fort Worth, v. 27, p. 509-540, 1992.

BURGUENO, J. et al. Genomic prediction of breeding values when modeling genotype × environment interaction using pedigree and dense molecular markers. **Crop Science**, Madison, v. 52, p. 707-719, 2012.

BURGUEÑO, J. et al. Modeling additive × environment and additive × additive × environment using genetic covariances of relatives of wheat genotypes. **Crop Science**, Madison, v. 43, p. 311-320, 2007.

BURGUEÑO, J. et al. Using factor analytic models for joining environments and genotypes without crossover genotype × environment interaction. **Crop Science**, Madison, v. 48, p. 1291-1305, 2008.

BUTLER, D. G. et al. **Mixed models for S language environments, ASReml-R reference manual**. Brisbane: QLD Departmentof Primary Industries and Fisheries, 2009. 145 p. (Training and Development Series, QE02001).

CAMPOS, G. de los; GIANOLA, D. Factor analysis models for structuring covariance matrices of additive genetic effects: a Bayesian implementation. **Genetics, Selection, Evolution**, Paris, v. 39, n. 5, p. 481-494, 2007.

CAO, Y. **A Bayesian approach to factor analysis via comparing posteriori and prior concentration**. 2010. 177 p. Thesis (Ph.D. in Department of Statistics)-University of Toronto, Toronto, 2010.

CAO, Y.; EVANS, M.; GUTTMAN, I. **Bayesian factor analysis via concentration**. Toronto: University of Toronto, 2010. 25 p. (Technical Report, 1003).

CHEN, M. H.; SHAO, Q. M. Monte Carlo estimation of bayesian credible and HPD intervals. **Journal of Computational and Graphical Statistics**, Alexandria, v. 8, n. 1, p. 69-92, July 1999.

CORNELIUS, P. L.; CROSSA, J. Prediction assessment of shrinkage estimators of multiplicative models for multi-environment cultivar trials. **Crop Science**, Madison, v. 39, p. 998-1009, 1999.

CORNELIUS, P. L.; CROSSA, J.; SEYEDSADR, M. S. Statistical tests and estimators of multiplicative models for genotype-by-environment interaction. In: KANG, M. S.; GAUCH, H. G. (Org.). **Genotype-by-environment interaction**. Boca Raton: CRC, 1996. p. 199-234.

CORNELIUS, P. L.; SEYEDSADR, M. S. Estimation of general linear-bilinear models for twoway tables. **Journal of Statistical Computation and Simulation**, New York, v. 58, p. 287-322, 1997.

COSTA, W. D. **Técnicas bayesianas para engenharia elétrica**. 2004. Disponível em: <http://www.cpdee.ufmg.br/~wadaed/Pesquisa/TecBayesianas.pdf>. Acesso em: 10 mar. 2016.

COTES, J. et al. A Bayesian approach for assessing the stability of genotypes. **Crop Science**, Madison, v. 46, p. 2654-2665, 2006.

CROSSA, J. From genotype × environment interaction to gene × environment interaction. **Current Genomics**, Beijing, v. 13, n. 3, p. 225-244, 2012.

CROSSA, J. Statistical analyses of multilocation trials. **Advances in Agronomy**, San Diego, v. 44, p. 55-85, 1990.

CROSSA, J.; CORNELIUS, L. Linear-bilinear models for the analysis of genotype environment interaction. In: KANG, M. S. (Ed.). **Quantitative genetics, genomics and plant breeding**. Oxford: CAB International, 2002. p. 305-322.

CROSSA, J. et al. Bayesian estimation of the additive main effects and multiplicative interaction model. **Crop Science**, Madison, v. 51, n. 4, p. 1458-1469, July 2011.

CROSSA, J. et al. Modeling genotype × environment interaction using additive genetic covariances of relatives for predicting breeding values of wheat genotypes. **Crop Science**, Madison, v. 46, p. 1722-1733, 2006.

CROSSA, J.; YANG, R. C.; CORNELIUS, P. L. Studying crossover genotype × environment interaction using linear-bilinear models and mixed models. **Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 9, p. 362-380, 2004.

DICICCIO, T. et al. Computing Bayes' factors by combining simulation and asymptotic approximations. **Journal of the American Statistical Association**, New York, v. 92, p. 903-915, 1997.

EBERHART, S. A.; RUSSELL, W. A. Stability parameters for comparing varieties. **Crop Science**, Madison, v. 6, p. 36-40, 1966.

EDWARDS, J. W.; JANNINK, J. L. Bayesian modeling of heterogeneous error and genotype environment interaction variances. **Crop Science**, Madison, v. 46, p. 820-833, 2006.

FINLAY, K. W.; WILKINSON, G. N. The analysis of adaptation in a plantbreeding programme. **Australian Journal of Agricultural Research**, Collingwood, v. 14, p. 742-754, 1963.

GABRIEL, K. R. Le biplot: outil d'exploration de données multidimensionelles. **Journal de la Societe Francaise de Statistique**, Paris, v. 143, p. 5-55, 2002.

GABRIEL, K. R. The biplot graphic display of matrices with application to principal component analysis. **Biometrika**, London, v. 58, p. 453-467, 1971.

GABRIEL, K. R. Least squares approximation of matrices by additive and multiplicative models. **Journal of the Royal Statistical Society: Series B**, London, v. 40, p. 186-196, 1978.

GAUCH, H. G. Model selections and validation for yield trials with interactions. **Biometrics**, Washington, v. 44, p. 705-715, 1988.

GAUCH, H. G. Statistical analysis of yield trials by AMMI and GGE. **Crop Science**, Madison, v. 46, p. 1488-1500, 2006.

GAUCH, H. G.; PIEPHO, H. P.; ANNICCHIARICO, P. Statistical analysis of yield trials by AMMI and GGE: further considerations. **Crop Science**, Madison, v. 48, p. 866-889, 2008.

GEISSER, S.; EDDY, W. F. A predictive approach to model selection. **Journal of the American Statistical Association**, New York, v. 74, p. 153-160, 1979.

GELFAND, A. E. Model determination using sampling-based methods. In: GILKS, W. R.; RICHARDSON, S.; SPIEGELHALTER, D. J. (Ed.). **Markov chain Monte Carlo inpractice**. London: Chapman & Hall, 1996. p. 145-161.

GELMAN, A. et al. **Bayesian data analysis**. 3rd ed. Boca Raton: Chapman and Hall; CRC, 2013. 675 p.

GELMAN, A.; MENG, X. L. Simulating normalizing constants: from importance sampling to bridge sampling to path sampling. **Statistical Science**, Hayward, v. 13, p. 163-185, 1998.

GEWEKE, J. F.; ZHOU, G. Measuring the pricing error of the arbitrage pricing Multilevel Factor Analysis theory. **Review of Financial Stududies**, Oxford, v. 9, p. 557-587, 1996.

GHOSH, J.; DUNSON, D. B. Default priors and efficient posterior computation in Bayesian factor analysis. **Journal of Computational and Graphical Statistics**, Alexandria, v. 18, n. 2, p. 306-320, June 2009.

GOLLOB, H. F. A statistical model which combines features of factoranalytic and analysis of variance. **Psychometrika**, Williamsburg, v. 33, p. 73-115, 1968.

GOSH, J. B.; DUNSON, D. Bayesian model selection in factor analytic models. In: DUNSON, D. B. (Ed.). **Random effect and latent variable model selection**. New York: J. Wiley, 2008. p. 151-163.

GREEN, P. J. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. **Biometrika**, London, v. 82, n. 4, p. 711-732, 1995.

HAYASHI, K.; SEN, P. K. Bayesian factor analysis: bias-corrected estimator of factor loadings in Bayesian factor analysis. **Educational and Psychological Measurement**, London, v. 62, n. 6, p. 944-959, 2002.

HAYASHI, K.; YUNG, Y. F. Standard errors for the class of orthomax-rotated factor pattern coefficients: some matrix results. **Psychometrika**, v. 64, p. 451-460, 1999.

HEIDELBERGER, P.; WELCH, P. D. Simulation run length control in the presence of an initial transient. **Operations Research**, Catonsville, v. 31, p. 1109-1144, 1983.

HILL, W. G. On selection among groups with heterogeneous variance. **Animal Production**, Edinburgh, v. 39, p. 473-477, 1984.

HOFF, P. D. Simulation of the matrix bingham-von mises-fisher distribution, with applications to multivariate and relational data. **Journal of Computational and Graphical Statistics**, Alexandria, v. 18, p. 438-456, 2009.

HOGAN, J. W.; TCHERNIS, R. Bayesian factor analysis for spatially correlated data, with application to summarizing area-Level material deprivation from census data. **Journal of the American Statistical Association**, New York, v. 99, p. 314-324, 2004.

HU, Z.; YANG, R. C. Improved statistical inference for graphical description and interpretation of genotype times environment interaction. **Crop Science**, Madison, v. 53, n. 6, p. 2400-2410, 2013a.

HU, Z.; YANG, R. C. A new distribution-free approach to constructing the confidence region for multiple parameters. **PLoS One**, San Francisco, v. 8, p. e81179, 2013b.

HUYNH, H. S.; FELDT, L. S. Conditions under which mean square ratios in repeated measurements designs have exact F distributions. **Journal of the American Statistical Association**, New York, v. 65, p. 1582-1589, 1970.

HUYNH, H. S.; FELDT, L. S. Estimation of the Box correction for degrees of freedom from sample data in randomized block and split-plot designs. **Journal of Educational Statistics**, Washington, v. 1, p. 69-82, 1976.

JARQUÍN, D. et al. A hierarchical Bayesian estimation model for multi-environment plant breeding trials in successive years. **Crop Science**, Madison, v. 56, p. 1-17, Aug. 2016.

JENNRICH, R. I.; SCHLUCHTER, M. D. Unbalanced repeated-measures models with structured covariance matrices. **Biometrics**, Washington, v. 42, p. 805-820, 1986.

JOHNSON, R. A.; WICHERN, D. W. **Applied multivariate statistical analysis**. Englewood Cliffs: Prentice-Hall, 2007. 767 p.

JOSSE, J. et al. Another look at Bayesian analysis of AMMI models for genotype-environment data. **Journal of Agricultural, Biological and Environmental Statistics**, Alexandria, v. 19, p. 240-257, 2014.

KARATZA, A. S. **Bayesian factor analysis:** implementation on schizotypal personality disorder data. 2006. 136 p. Thesis (Ph.D. in Statistics)-Athens University of Economics and Business, Athens, 2006.

KAUFMAN, G. M.; PRESS, S. J. **Bayesian factor analysis**. Chicago: University of Chicago, 1973. 30 p. (Technical Report, 7322).

KELLY, A. M. et al. The accuracy of varietal selection using factor analytic models for multi-environment plant breeding trials. **Crop Science**, Madison, v. 47, p. 1063-1070, 2007.

KELLY, A. M. et al. Estimation in a multiplicative mixed model involving a genetic relationship matrix. **Genetics Selection Evolution**, Paris, v. 41, n. 1, p. 1-9, Apr. 2009.

KESELMAN, H. J. et al. A comparison of two approaches for selecting covariance structures in the analysis of repeated measurements. **Communications in Statistics - Simulation and Computation**, Ontario, v. 27, p. 591-604, 1998.

LEE, S. E. **Robustness of Bayesian factor analysis estimates**. 1994. Thesis (Ph.D. in Statistics)-University of California, Riverside, 1994.

LEE, S. Y. **Structural equation modeling:** a Bayesian approach. Chichester: J. Wiley, 2007. 458 p.

LEE, S. Y.; SONG, X. Y. Bayesian selection on the number of factors in a factor analysis model. **Behaviormetrika**, New York, v. 29, p. 23-40, 2002.

LITTELL, R. C.; PENDERGAST, J.; NATARAJAN, R. Modelling covariance structure in the analysis of repeated measures data. **Statistics in Medicine**, New York, v. 19, p. 1793-1819, 2000.

LIU, G. **Bayesian computation for general linear-bilinear models**. 2001. 150 p. Dissertation (Ph.D. in Statistics)-University of Kentucky, Lexington, 2001.

LOPES, H. F.; CARVALHO, C. M. Factor stochastic volatility with time varying loadings and Markov switching regimes. **Journal of Statistical Planning and Inference**, New York, v. 137, p. 3082-3091, 2007.

LOPES, H. F.; WEST, M. Bayesian model assessment in factor analysis. **Statistica Sinica**, Taipei, v. 14, n. 1, p. 41-67, 2004.

MALAKOFF, D. M. Bayes offers “New” way to make sense of numbers. **Science**, New York, v. 286, p. 1460-1464, 1999.

MANDEL, J. The partitioning of interaction in analysis of variance. **Journal of Research of the National Bureau Standards, Series B**, Washington, v. 73, p. 309-328, 1969.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. London: Academic, 1979. 521 p.

MARDIA, K. V.; KENT, J. T.; BIBBY, J. M. **Multivariate analysis**. London: Academic, 1988. 512 p.

MELO, W. M. C. et al. Prediction of maize single cross hybrids using the total effects of associated markers approach assessed by cross-validation and regional trials. **Scientific World Journal**, London, v. 2014, p. 1-9, 2014.

MENG, X. L.; WONG, W. Simulating ratios of normalizing constants via a simple Identity: a theoretical exploration. **Statistica Sinica**, Taipei, v. 6, p. 831-860, 1996.

MEYER, K. Factor-analytic models for genotype x environment type problems and structured covariance matrices. **Genetics Selection Evolution**, Paris, v. 41, p. 21, Jan. 2009.

MEYER, K.; KIRKPATRICK, M. Perils of parsimony: properties of reduced rank estimates of genetic covariances. **Genetics**, New York, v. 108, p. 1153-1166, 2008.

NEUHAUS, J. O.; WRIGLEY, C. F. Further investigations in factor estimation. **Proceedings of the Royal Society of Edinburgh**, Edinburgh, v. 61, p. 176-185, 1941.

NUVUNGA, J. J. et al. Factor analysis using mixed models of multi-environment trials with different levels of unbalancing. **Genetics and Molecular Research**, Ribeirão Preto, v. 14, p. 14262-14278, Nov. 2015.

OAKEY, H. et al. Genomic selection in multi-environment crop trials. **Genes**, Bethesda, v. 6, n. 5, p. 1313-1326, May 2016.

OAKEY, H. et al. Joint modeling of additive and non-additive (genetic line) effects in multi-environment trials. **Theoretical and Applied Genetics**, Berlin, v. 114, p. 1319-1332, 2007.

OLIVEIRA, L. A. de et al. Bayesian GGE biplot models applied to maize multi-environments trials. **Genetics and Molecular Research**, Ribeirão Preto, v. 15, n. 2, p. 1-21, 2016.

OLIVEIRA, L. A. de et al. Credible intervals for scores in the AMMI with random effects for genotype. **Crop Science**, Madison, v. 55, p. 465-476, 2015.

OOMS, J. C. L. **The highest posterior density posterior prior for Bayesianmodel selection**. 2009. Disponível em: <http://igiturarchive.library.uu.nl/studenttheses/0708-202210/thesis-Jeroen_Ooms.pdf>. Acesso em: 12 nov. 2015.

ORELLANA, M. A. **Bayesian prediction of crop performance modeling genotype by environment interaction with heterogeneous variances**. 2012. Paper 12740. Disponível em: <http://lib.dr.iastate.edu/etd/12740>. Acesso em: 10 out. 2016.

ORELLANA, M. A.; EDWARDS, J.; CARRIQUIRY, A. Heterogeneous variances in multi-environment yield trials for corn hybrids. **Crop Science**, Madison, v. 54, p. 1048-1056, 2014.

PEREZ-ELIZALDE, S.; JARQUIN, D.; CROSSA, J. A general Bayesian estimation method of linear-bilinear models applied to plant breeding trials with genotype × environment interaction. **Journal of Agricultural, Biological and Environmental Statistics**, Alexandria, v. 17, n. 1, p. 15-37, 2011.

PIEPHO, H. P. Analysing genotype-environment data by mixed models with multiplicative terms. **Biometrics**, Washington, v. 53, p. 761-767, 1997.

PIEPHO, H. P. Empirical best linear unbiased prediction in cultivar trials using factor-analytic variance-covariance structures. **Theoretical Applied Genetics**, Berlin, v. 97, p. 195-201, 1998.

POLASEK, W. **Factor analysis and outliers:** a Bayesian approach. Basel: University of Basel, 1997. 26 p. Discussion paper.

PRESS, S. J. **Applied multivariate analysis:** using Bayesian and frequentist methods of inference. Melbourne: Krieger, 1972. 500 p.

PRESS, S. J. **Applied multivariate analysis:** using bayesian and frequentist methods of inference. Melbourne: Krieger, 1985. 704 p.

PRESS, S. J. **Bayesian statistics:** principles, models, and applications. New York: J. Wiley, 1989. 256 p.

PRESS, S. J.; SHIGEMASU, K. Bayesian inference in factor analysis. In: GLESER, L. J. et al. (Ed.). **Contributions to probability and statistics:** essays in Honor of Ingram Olkin. New York: Springer-Verlag, 1989. p. 271-287.

PRESS, S. J.; SHIGEMASU, K. **Bayesian inference in factor analysis revised**. Riverside: University of California, 1997. 29 p. (Technical Report, 243).

PRESS, S. J.; SHIGEMASU, K. A note on choosing the number of factors. **Communications in Statistics - Theory and Methods**, Toronto, v. 28, p. 1653-1670, 1999.

R CORE TEAM. **R:** a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2016. Disponível em: <https://www.R-project.org/>. Acesso em: 10 out. 2016.

RAFTERY, A. E. et al. Estimating the integrated likelihood via posterior simulation using the harmonic mean identity. **Bayesian Statistics**, Kyoto, v. 8, p. 1-45, 2007.

RAFTERY, A. E.; LEWIS, S. M. How many iterations in the Gibbs sampler? In: _____. **Bayesian statistics 4**. Oxford: Oxford University Press, 1992. p. 763-773.

RAFTERY, A. E.; LEWIS, S. M. The number of iterations, convergence diagnostics and generic metropolis algorithms. In: GILKS, W. R.; SPIEGELHALTER, D. J. (Ed.). **Practical Markov Chain Monte Carlo**. London: Chapman and Hall, 1995. p. 115-130.

RONNEGARD, L. et al. Genetic heterogeneity of residual variance: estimation of variance components using double hierarchical generalized linear models. **Genetics Selection Evolution**, Paris, v. 42, n. 1, p. 1-10, 2010.

ROWE, D. B. **Correlated Bayesian factor analysis**. 1998. 155 p. Thesis (Ph.D. in Statistics)-University of California, Riverside, 1998.

ROWE, D. B. **Multivariate Bayesian statistics:** models for source separation and signal unmixing. Boca Raton: Chapmanand Hall; CRC, 2002. 352 p.

ROWE, D. B. **Multivariate Bayesian statistics:** models for source separation and signal unmixing. Boca Raton: CRC, 2003. 323 p.

RUNCIE, D. E.; MUKHERJEE, S. Dissecting high-dimensional phenotypes with Bayesian sparse factor analysis of genetic covariance matrices. **Genetics**, Austin, v. 194, n. 3, p. 753-776, July 2013.

SCHAEFFER, L. R. Multiple-country comparison of dairy sires. **Journal of Dairy Science**, Champaign, v. 77, p. 2671-2678, 1994.

SCHWARTZ, G. Estimating the dimension of a model. **Annals of Statistics**, Hayward, v. 6, p. 461-464, 1978.

SILVA, C. P. da et al. A Bayesian Shrinkage approach for AMMI Models. **PLoS One**, San Francisco, v. 10, n. 7, p. e0131414, 2015.

SMITH, A. et al. Multiplicative mixed models for the analysis of sensory evaluation data. **Food Quality and Preference**, Barking, v. 14, n. 5/6, p. 387-395, 2003.

SMITH, A. B.; CULLIS, B. R.; THOMPSON, R. Analysing variety by environment data using multiplicative mixed models and adjustments for spatial fi eld trend. **Biometrics**, Washington, v. 57, p. 1138-1147, 2001.

SMITH, A. B.; CULLIS, B. R.; THOMPSON, R. The analysis of crop cultivar breeding and evaluation trials: an overview of current mixed model approaches. **Journal of Agricultural Science**, Cambridge, v. 143, p. 1-14, 2005.

SMITH, A. B. et al. Exploring variety-environment data using random effects AMMI models with adjustments for spatial field trend: part II, applications. In: KANG, M. (Ed.). **Quantitative genetics, genomics and plant breeding**. London: CABI, 2002. p. 337-352.

SMITH, A. B. et al. Factor analytic mixed models for the provision of grower information from national crop variety testing programs. **Theoretical Applied Genetics**, Berlin, v. 128, n. 1, p. 55-72, Jan. 2015.

SONGGUI, W.; SUJU, Y. A new estimate of the parameters in linear mixed models. **Science in China. (Series A)**, Beijing, v. 45, n. 10, p. 1301-1311, Oct. 2002.

STEFANOVA, K.; BUIRCHELL, B. Multiplicative mixed models for genetic gain assessment in lupin breeding. **Crop Science**, Madison, v. 50, p. 880-891, May/June 2010.

SUN, L. et al. Bayesian methods for variance component models. **Journal of the American Statistical Association**, New York, v. 91, n. 434, p. 743-752, 1996.

TANNER, M. A. **Tools for statistical inference:** methods for the exploration of posterior distributions and likelihood functions. 2nd ed. New York: Springer Verlag, 1993. 208 p.

THEOBALD, C. M.; TALBOT, M.; NABUGOOMU, F. A Bayesian approach to regional and local-area prediction from crop variety trials. **Journal of Agricultural, Biological, and Environmental Statistics**, Alexandria, v. 7, p. 403-419, 2002.

THURSTONE, L. L. A new rotational method in factor analysis. **Psychometrika**, Williamsburg, v. 3, p. 199-218, 1938.

TYRISEVÄ, A. M. et al. Principal component approach in variance component estimation for international sire evaluation. **Genetic Selection Evolution**, Paris, v. 24, n. 43, p. 21, May 2011.

VIELE, K.; SRINIVASAN, C. Parsimonious estimation of multiplicative interaction in analysis of variance using Kullback-Leibler information. **Journal of Statistical Planning and Inference**, Amsterdam, v. 84, p. 201-219, 2000.

WEST, E. Gene expression predictors of breast cancer outcomes. **The Lancet**, London, v. 361, n. 9396, p. 1590-1596, 2003.

WILLIAMS, E. J. The interpretation of interactions in factorial experiments. **Biometrika**, Washington, v. 39, p. 65-81, 1952.

WOLFINGER, R. Covariance structure selection in general mixed models. **Communications in Statistics - Simulation**, Ontario, v. 22, n. 4, p. 1079-1106, 1993.

XAVIER, L. H. **Modelos univariado e multivariado para análise de medidas repetidas e verificação da acurácia do modelo univariado por meio de simulação**. 2000. 91 p. Dissertação (Mestrado em Estatística Experimental)-Escola Superior de Agricultura "Luiz de Queiroz", Piracicaba, 2000.

YAN, W. Comment on “Biplot Analysis of Genotype × environment interaction: proceed with caution”. **Crop Science**, Madison, v. 50, p. 1121-1123, 2010.

YAN, W. et al. Cultivar evaluation and mega-environment investigation based on the GGE biplot. **Crop Science**, Madison, v. 40, n. 3, p. 597-605, May/June 2000.

YAN, W. et al. GGE biplot vs. AMMI analysis of genotype-by-environment data. **Crop Science**, Madison, v. 47, p. 643-655, 2007.

YANG, R. et al. Biplot analysis of Genotype × environment interaction: proceed with caution. **Crop Science**, Madison, v 49, p. 1564-1576, 2009.

YATES, F.; COCHRAN, W. G. The analysis of groups of experiments. **Journal of Agricultural Science**, Cambridge, v. 28, p. 556-580, 1938.

ZELLNER, A.; CHUNG-KI, M. Bayesian analysis, model selection and prediction. In: _____. **Physics and probability:** essays in Honor of Edwin T. Jaynes. Cambridge: Cambridge University Press, 1993. p. 195-206.

ZHANG, N. L.; KOCKA, T. Effective dimensions of hierarchical latent class models. **Journal of Artificial Intelligence Research**, Palo Alto, v. 21, p. 1-17, 2004.

APÊNDICE A-demonstração da expansão matricial do modelo

Expansão do modelo da fórmula

Considerando o modelo para casela com dois ambientes e 3 genótipos e três blocos.

O modelo misto multivariado padrão:

$$\mathbf{y = X\beta + Zu + \varepsilon} \qquad (1)$$

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{311} \\ y_{312} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \\ y_{321} \\ y_{322} \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} u_{11} \\ u_{12} \\ u_{13} \\ u_{21} \\ u_{22} \\ u_{32} \end{bmatrix} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{311} \\ \varepsilon_{312} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{221} \\ \varepsilon_{222} \\ \varepsilon_{321} \\ \varepsilon_{322} \end{bmatrix}$$

$y_{111} = \mu + \beta_1 + u_{11} + \varepsilon_{111}$ e $y_{112} = \mu + \beta_1 + u_{12} + \varepsilon_{112}$

Usando o modelo proposto por Smith et al. 2001

$$\mathbf{Zu} = \left[\mathbf{Z_1} \ldots \mathbf{Z}_m\right] \begin{bmatrix} u_1 \\ \vdots \\ u_m \end{bmatrix} = \mathbf{Z_1}u_1 + \cdots + \mathbf{Z}_m u_m$$

$$\mathbf{u = (\Gamma_m \otimes I_p)f + \delta}$$

Então temos um novo modelo

$$\mathbf{y = X\beta + Z(\Gamma_m \otimes I_p)f + Z\delta + \varepsilon} \qquad (2)$$

$$\mathbf{\Gamma}_m = \begin{bmatrix} \tau_{11} & \tau_{12} \\ \tau_{21} & \tau_{22} \end{bmatrix}$$

$$\left(\boldsymbol{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}}\right)=\begin{bmatrix}\tau_{11} & \tau_{12}\\ \tau_{21} & \tau_{22}\end{bmatrix}\otimes\begin{bmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}\tau_{11} & 0 & 0 & \tau_{12} & 0 & 0\\ 0 & \tau_{11} & 0 & 0 & \tau_{12} & 0\\ 0 & 0 & \tau_{11} & 0 & 0 & \tau_{12}\\ \tau_{21} & 0 & 0 & \tau_{22} & 0 & 0\\ 0 & \tau_{21} & 0 & 0 & \tau_{22} & 0\\ 0 & 0 & \tau_{21} & 0 & 0 & \tau_{22}\end{bmatrix}$$

Considerando o modelo full temos

$m=2; p=3$

$$(\boldsymbol{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}})\mathbf{f}=\begin{bmatrix}\tau_{11} & 0 & 0 & \tau_{12} & 0 & 0\\ 0 & \tau_{11} & 0 & 0 & \tau_{12} & 0\\ 0 & 0 & \tau_{11} & 0 & 0 & \tau_{12}\\ \tau_{21} & 0 & 0 & \tau_{22} & 0 & 0\\ 0 & \tau_{21} & 0 & 0 & \tau_{22} & 0\\ 0 & 0 & \tau_{21} & 0 & 0 & \tau_{22}\end{bmatrix}_{6\times 6}\begin{bmatrix}f_{11}\\ f_{12}\\ f_{13}\\ f_{21}\\ f_{22}\\ f_{23}\end{bmatrix}_{6\times 1}=\begin{bmatrix}\tau_{11}f_{11}+\tau_{12}f_{21}\\ \tau_{11}f_{21}+\tau_{12}f_{22}\\ \tau_{11}f_{13}+\tau_{12}f_{23}\\ \tau_{21}f_{11}+\tau_{22}f_{21}\\ \tau_{21}f_{12}+\tau_{22}f_{22}\\ \tau_{21}f_{13}+\tau_{22}f_{22}\end{bmatrix}=\begin{bmatrix}u_{11}\\ u_{12}\\ u_{13}\\ u_{21}\\ u_{22}\\ u_{32}\end{bmatrix}$$

$$\mathbf{Z}(\boldsymbol{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}})\mathbf{f}=\begin{bmatrix}1 & 0 & 0 & 0 & 0 & 0\\ 1 & 0 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 0 & 0 & 1\\ 0 & 0 & 0 & 0 & 0 & 1\end{bmatrix}_{12\times 6}\begin{bmatrix}\tau_{11}f_{11}+\tau_{12}f_{21}\\ \tau_{11}f_{21}+\tau_{12}f_{22}\\ \tau_{11}f_{13}+\tau_{12}f_{23}\\ \tau_{21}f_{11}+\tau_{22}f_{21}\\ \tau_{21}f_{12}+\tau_{22}f_{22}\\ \tau_{21}f_{13}+\tau_{22}f_{22}\end{bmatrix}_{6\times 1}=\begin{bmatrix}\tau_{11}f_{11}+\tau_{12}f_{21}\\ \tau_{11}f_{11}+\tau_{12}f_{21}\\ \tau_{11}f_{12}+\tau_{12}f_{22}\\ \tau_{11}f_{12}+\tau_{12}f_{22}\\ \tau_{11}f_{13}+\tau_{12}f_{23}\\ \tau_{11}f_{13}+\tau_{12}f_{23}\\ \tau_{21}f_{11}+\tau_{22}f_{21}\\ \tau_{21}f_{11}+\tau_{22}f_{21}\\ \tau_{21}f_{12}+\tau_{22}f_{22}\\ \tau_{21}f_{12}+\tau_{22}f_{22}\\ \tau_{21}f_{13}+\tau_{22}f_{23}\\ \tau_{21}f_{13}+\tau_{22}f_{23}\end{bmatrix}_{12\times 1}$$

$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{311} \\ y_{312} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \\ y_{321} \\ y_{322} \end{bmatrix}_{12\times1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \tau_{11} f_{11} + \tau_{12} f_{21} \\ \tau_{11} f_{11} + \tau_{12} f_{21} \\ \tau_{11} f_{12} + \tau_{12} f_{22} \\ \tau_{11} f_{12} + \tau_{12} f_{22} \\ \tau_{11} f_{13} + \tau_{12} f_{23} \\ \tau_{11} f_{13} + \tau_{12} f_{23} \\ \tau_{21} f_{11} + \tau_{22} f_{21} \\ \tau_{21} f_{11} + \tau_{22} f_{21} \\ \tau_{21} f_{12} + \tau_{22} f_{22} \\ \tau_{21} f_{12} + \tau_{22} f_{22} \\ \tau_{21} f_{13} + \tau_{22} f_{23} \\ \tau_{21} f_{13} + \tau_{22} f_{23} \end{bmatrix}_{12\times1} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{311} \\ \varepsilon_{312} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{221} \\ \varepsilon_{222} \\ \varepsilon_{321} \\ \varepsilon_{322} \end{bmatrix}_{12\times1}$$

Considerando a decomposição espectral

$$\hat{\mathbf{\Gamma}}_k = V\Lambda^{\frac{1}{2}}$$

Em que: $\mathbf{\Lambda} = diag(\lambda_1,...\lambda_k)$ e $\mathbf{V} = (\boldsymbol{\alpha}_1,...\boldsymbol{\alpha}_k)$.

$$\text{E } \tau_{kk} = \lambda_k a_{pk}$$

$$\mathbf{Z}(\mathbf{\Gamma}_{\mathrm{m}} \otimes \mathbf{I}_{\mathrm{p}})\mathbf{f} = \sum_{k=1}^{m} \lambda_k \mathbf{diag}\left(\mathbf{X}_2\boldsymbol{\alpha}_\mathbf{k}\right)\mathbf{Z}_1\mathbf{f}_\mathbf{k}$$

O modelo (2) pode ser reescrito:

$$\mathbf{y} = \mathbf{X}_1\boldsymbol{\beta} + \sum_{k=1}^{m} \lambda_k \mathbf{diag}(\mathbf{X}_2\boldsymbol{\alpha}_\mathbf{k})\mathbf{Z}_1\mathbf{f}_\mathbf{k} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

$$\begin{array}{c} \\ \mathbf{G1} \\ \\ \mathbf{G2} \\ \\ \mathbf{G3} \\ \\ \end{array}\begin{array}{c} \begin{array}{cc} \mathbf{A1} & \mathbf{A2} \end{array} \\ \begin{bmatrix} y_{111} & y_{121} \\ y_{112} & y_{122} \\ y_{211} & y_{221} \\ y_{212} & y_{222} \\ y_{311} & y_{321} \\ y_{312} & y_{322} \end{bmatrix} \end{array} \qquad \mathbf{X}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \qquad \mathbf{Z}_1 = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\boldsymbol{\alpha} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}; \qquad \boldsymbol{\alpha}_1 = \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix}; \quad \boldsymbol{\alpha}_2 = \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix}$$

$$\mathbf{f} = \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \\ f_{21} \\ f_{22} \\ f_{23} \end{bmatrix}; \quad \mathbf{f}_1 = \begin{bmatrix} f_{11} \\ f_{12} \\ f_{13} \end{bmatrix}; \quad \mathbf{f}_2 = \begin{bmatrix} f_{21} \\ f_{22} \\ f_{23} \end{bmatrix}$$

$$\mathbf{X}_2\boldsymbol{\alpha}_1 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{11} \\ a_{21} \end{bmatrix} = \begin{bmatrix} a_{11} \\ a_{11} \\ a_{11} \\ a_{11} \\ a_{11} \\ a_{11} \\ a_{21} \\ a_{21} \\ a_{21} \\ a_{21} \\ a_{21} \\ a_{21} \end{bmatrix} \qquad \mathbf{Z}_1\mathbf{f}_1 = \begin{bmatrix} f_{11} \\ f_{11} \\ f_{12} \\ f_{12} \\ f_{13} \\ f_{13} \\ f_{11} \\ f_{11} \\ f_{12} \\ f_{12} \\ f_{13} \\ f_{13} \end{bmatrix}$$

$$\lambda_1 diag\left(\mathbf{X}_2\boldsymbol{\alpha}_1\right)=\begin{bmatrix}\lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & \lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & \lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & \lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & \lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{11} & 0 & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21} & 0 & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21} & 0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21} & 0 & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21} & 0 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21} & 0\\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_1 a_{21}\end{bmatrix}$$

$$\lambda_1 diag(\mathbf{X}_2\boldsymbol{\alpha}_1)\mathbf{Z}_1\mathbf{f}_1=\begin{bmatrix}\lambda_1 a_{11} f_{11}\\ \lambda_1 a_{11} f_{11}\\ \lambda_1 a_{11} f_{12}\\ \lambda_1 a_{11} f_{12}\\ \lambda_1 a_{11} f_{13}\\ \lambda_1 a_{11} f_{13}\\ \lambda_1 a_{21} f_{11}\\ \lambda_1 a_{21} f_{11}\\ \lambda_1 a_{21} f_{12}\\ \lambda_1 a_{21} f_{12}\\ \lambda_1 a_{21} f_{13}\\ \lambda_1 a_{21} f_{13}\end{bmatrix}$$

$$\boldsymbol{\alpha}=\begin{bmatrix}a_{11} & a_{12}\\ a_{21} & a_{22}\end{bmatrix};\qquad \boldsymbol{\alpha}_1=\begin{bmatrix}a_{11}\\ a_{21}\end{bmatrix};\quad \boldsymbol{\alpha}_2=\begin{bmatrix}a_{12}\\ a_{22}\end{bmatrix}$$

$$\mathbf{f}=\begin{bmatrix}f_{11}\\ f_{12}\\ f_{13}\\ f_{21}\\ f_{22}\\ f_{23}\end{bmatrix};\quad \mathbf{f}_1=\begin{bmatrix}f_{11}\\ f_{12}\\ f_{13}\end{bmatrix};\quad \mathbf{f}_2=\begin{bmatrix}f_{21}\\ f_{22}\\ f_{23}\end{bmatrix}$$

$$\mathbf{X}_2\boldsymbol{\alpha}_2 = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a_{12} \\ a_{22} \end{bmatrix} = \begin{bmatrix} a_{12} \\ a_{12} \\ a_{12} \\ a_{12} \\ a_{12} \\ a_{12} \\ a_{22} \\ a_{22} \\ a_{22} \\ a_{22} \\ a_{22} \\ a_{22} \end{bmatrix} \qquad \mathbf{Z}_1\mathbf{f}_2 = \begin{bmatrix} f_{21} \\ f_{21} \\ f_{22} \\ f_{22} \\ f_{23} \\ f_{23} \\ f_{21} \\ f_{21} \\ f_{22} \\ f_{12} \\ f_{23} \\ f_{23} \end{bmatrix}$$

$$\lambda_2 diag(\mathbf{X}_2\boldsymbol{\alpha}_2) = \begin{bmatrix} \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{12} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 a_{22} \end{bmatrix}$$

$$\lambda_2 diag(\mathbf{X_2\alpha}_2)\mathbf{Z}_1\mathbf{f}_2 = \begin{bmatrix} \lambda_2 a_{12} f_{21} \\ \lambda_2 a_{12} f_{21} \\ \lambda_2 a_{12} f_{22} \\ \lambda_2 a_{12} f_{22} \\ \lambda_2 a_{12} f_{23} \\ \lambda_2 a_{12} f_{23} \\ \lambda_2 a_{22} f_{21} \\ \lambda_2 a_{22} f_{21} \\ \lambda_2 a_{22} f_{22} \\ \lambda_2 a_{22} f_{22} \\ \lambda_2 a_{22} f_{23} \\ \lambda_2 a_2 f_{23} \end{bmatrix}$$

$$\sum_{k=1}^{m} \lambda_k \mathbf{diag}(\mathbf{X_2\alpha_k})\mathbf{Z}_1\mathbf{f_k} = \lambda_1 diag(\mathbf{X_2\alpha}_1)\mathbf{Z}_1\mathbf{f}_1 + \lambda_2 diag(\mathbf{X_2\alpha}_2)\mathbf{Z}_1\mathbf{f}_2 =$$

$$= \begin{bmatrix} \lambda_1 a_{11} f_{11} \\ \lambda_1 a_{11} f_{11} \\ \lambda_1 a_{11} f_{12} \\ \lambda_1 a_{11} f_{12} \\ \lambda_1 a_{11} f_{13} \\ \lambda_1 a_{11} f_{13} \\ \lambda_1 a_{21} f_{11} \\ \lambda_1 a_{21} f_{11} \\ \lambda_1 a_{21} f_{12} \\ \lambda_1 a_{21} f_{12} \\ \lambda_1 a_{21} f_{13} \\ \lambda_1 a_{21} f_{13} \end{bmatrix} + \begin{bmatrix} \lambda_2 a_{12} f_{21} \\ \lambda_2 a_{12} f_{21} \\ \lambda_2 a_{12} f_{22} \\ \lambda_2 a_{12} f_{22} \\ \lambda_2 a_{12} f_{23} \\ \lambda_2 a_{12} f_{23} \\ \lambda_2 a_{22} f_{21} \\ \lambda_2 a_{22} f_{21} \\ \lambda_2 a_{22} f_{22} \\ \lambda_2 a_{22} f_{22} \\ \lambda_2 a_{22} f_{23} \\ \lambda_2 a_2 f_{23} \end{bmatrix} = \begin{bmatrix} \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \end{bmatrix}$$

Como podemos ver:

Usando componentes principais de e decomposição espectral e propriedades adequadas chegamos ao mesmo modelo full

$$\begin{bmatrix} \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \end{bmatrix}_{12\times 1} = \begin{bmatrix} \tau_{11} f_{11} + \tau_{12} f_{21} \\ \tau_{11} f_{11} + \tau_{12} f_{21} \\ \tau_{11} f_{12} + \tau_{12} f_{22} \\ \tau_{11} f_{12} + \tau_{12} f_{22} \\ \tau_{11} f_{13} + \tau_{12} f_{23} \\ \tau_{11} f_{13} + \tau_{12} f_{23} \\ \tau_{21} f_{11} + \tau_{22} f_{21} \\ \tau_{21} f_{11} + \tau_{22} f_{21} \\ \tau_{21} f_{12} + \tau_{22} f_{22} \\ \tau_{21} f_{12} + \tau_{22} f_{22} \\ \tau_{21} f_{13} + \tau_{22} f_{23} \\ \tau_{21} f_{13} + \tau_{22} f_{23} \end{bmatrix}_{12\times 1}$$

Logo a expansão do modelo na fórmula é dada por:

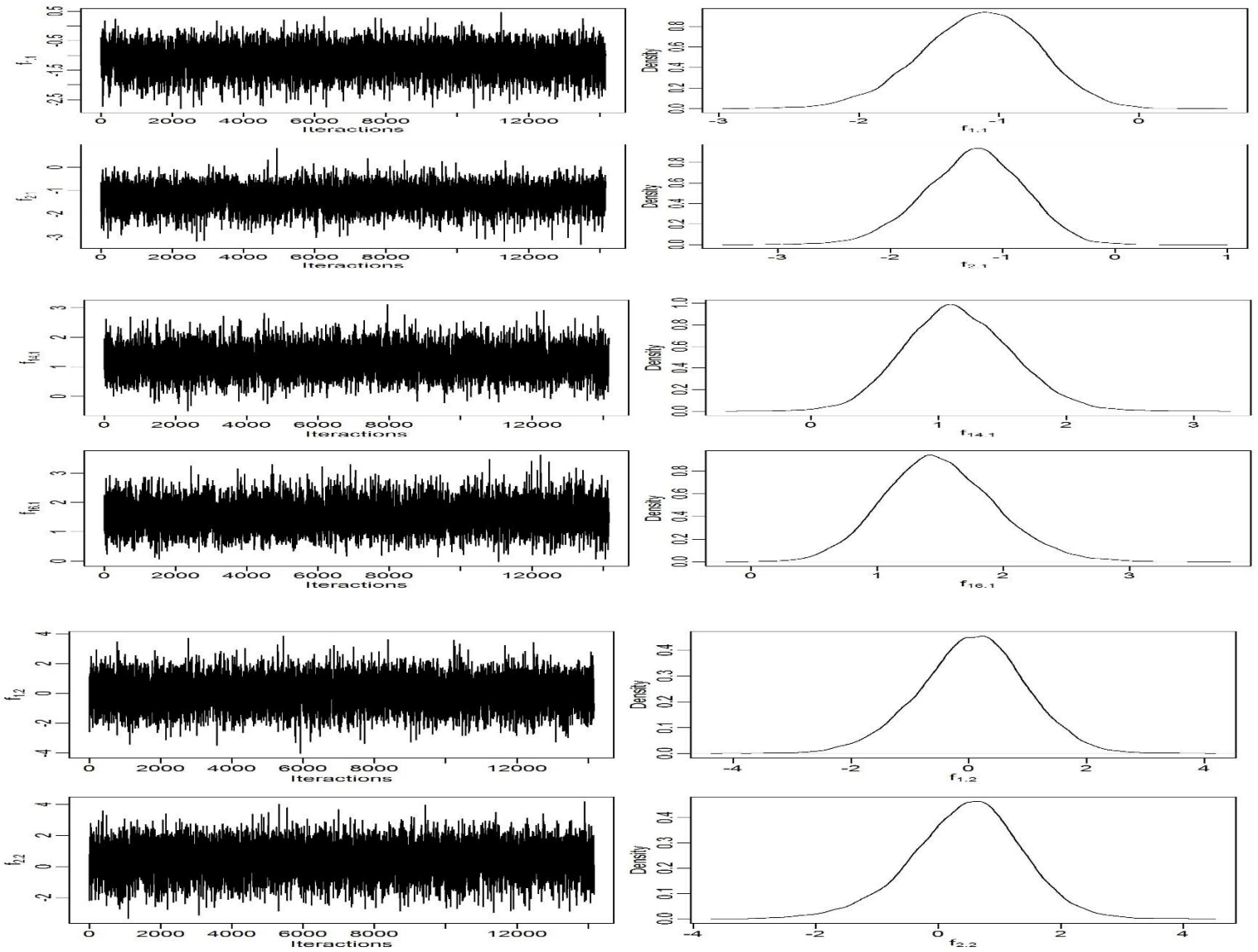
$$\mathbf{y} = \mathbf{X_1}\boldsymbol{\beta} + \sum_{k=1}^{m} \lambda_k \mathbf{diag}\left(\mathbf{X_2}\boldsymbol{\alpha_k}\right)\mathbf{Z_1}\mathbf{f_k} + \mathbf{Z}\boldsymbol{\delta} + \boldsymbol{\varepsilon}$$

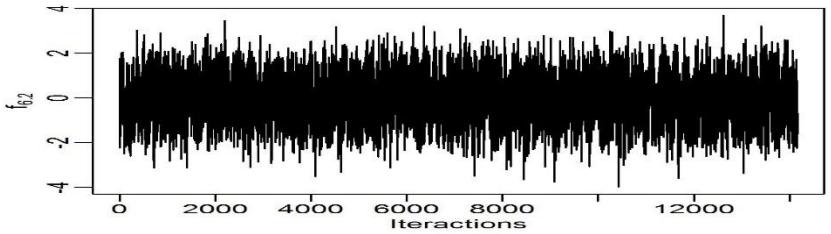
$$\begin{bmatrix} y_{111} \\ y_{112} \\ y_{211} \\ y_{212} \\ y_{311} \\ y_{312} \\ y_{121} \\ y_{122} \\ y_{221} \\ y_{222} \\ y_{321} \\ y_{322} \end{bmatrix}_{12\times 1} = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \beta_1 \\ \beta_2 \end{bmatrix} + \begin{bmatrix} \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{11} + \lambda_2 a_{12} f_{21} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{12} + \lambda_2 a_{12} f_{22} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{11} f_{13} + \lambda_2 a_{12} f_{23} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{11} + \lambda_2 a_{22} f_{21} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{12} + \lambda_2 a_{22} f_{22} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \\ \lambda_1 a_{21} f_{13} + \lambda_2 a_{22} f_{23} \end{bmatrix}_{12\times 1} + \begin{bmatrix} \varepsilon_{111} \\ \varepsilon_{112} \\ \varepsilon_{211} \\ \varepsilon_{212} \\ \varepsilon_{311} \\ \varepsilon_{312} \\ \varepsilon_{121} \\ \varepsilon_{122} \\ \varepsilon_{221} \\ \varepsilon_{222} \\ \varepsilon_{321} \\ \varepsilon_{322} \end{bmatrix}_{12\times 1}$$

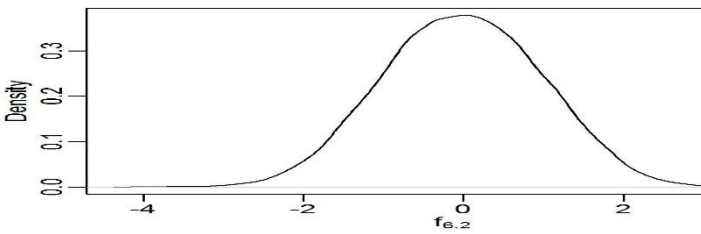
APÊNDICE B-traços de cadeias MCMC e densidades a posteriori

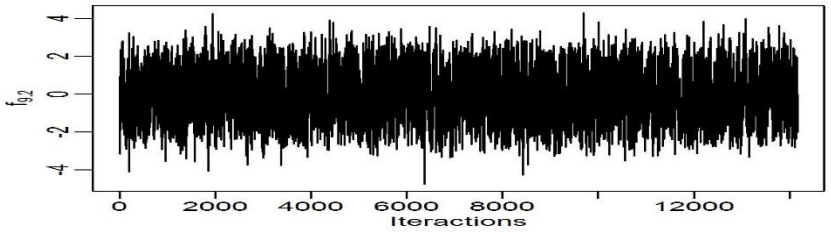
A seguir são apresentadas figuras referentes aos traços das cadeias MCMC para os k parâmetros. São apresentadas ainda os traços das cadeias e densidades, a posteriori, para os efeitos de genótipos e variância residual dos dados da simulação, cujos valores foram positivos e as densidades para as, bem como as densidades para os respectivos escores (genotípicos e ambientais), cujas regiões de credibilidade bivariadas, a 95% de credibilidade.

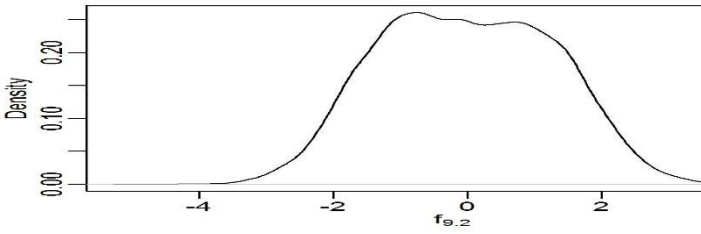
Figura 16- Traços das cadeias geradas pelo método MCMC e densidades a posteriori estimadas para os efeitos principais dos genótipos: G1, G2, G6,G9,G14, G16,G18 e G20

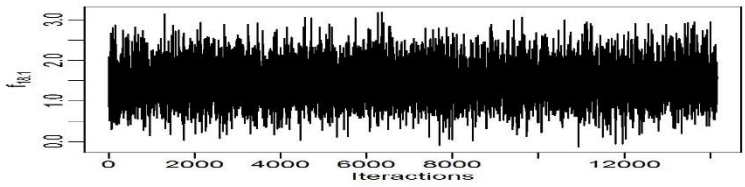


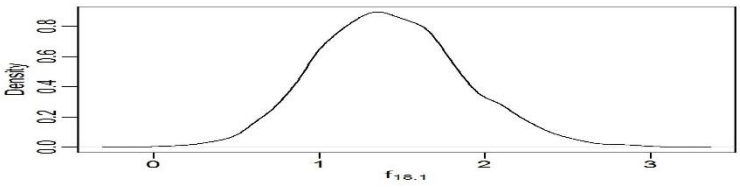

$f_{6.2}$
Iteractions
0 2000 4000 6000 8000 12000
-4 -2 0 2 4

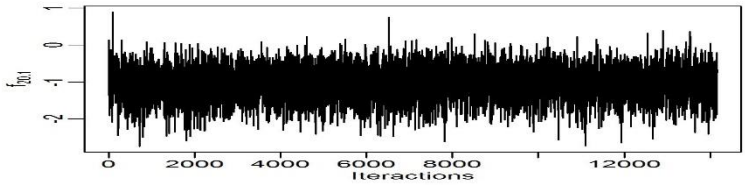

Density
$f_{6.2}$
0.0 0.1 0.2 0.3
-4 -2 0 2

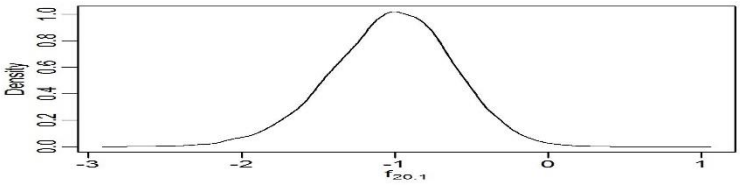

$f_{9.2}$
Iteractions
0 2000 4000 6000 8000 12000
-4 -2 0 2 4

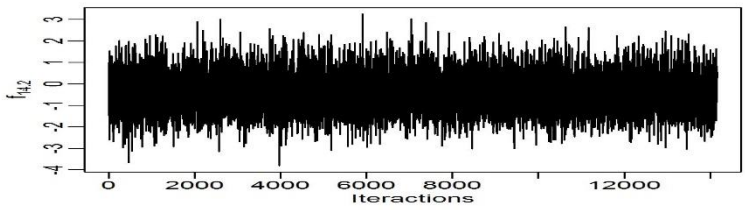

Density
$f_{9.2}$
0.00 0.10 0.20
-4 -2 0 2

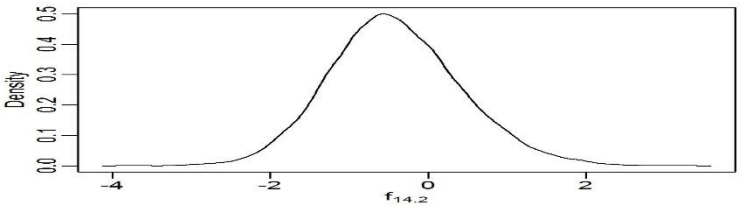

$f_{18.1}$
Iteractions
0 2000 4000 6000 8000 12000
0.0 1.0 2.0 3.0

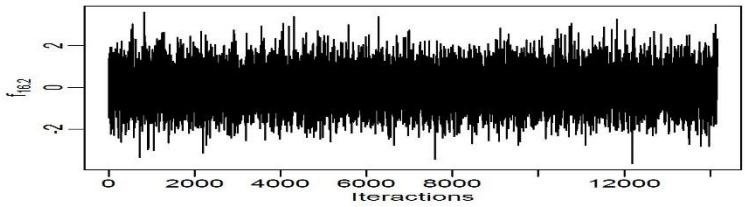

Density
$f_{18.1}$
0.0 0.2 0.4 0.6 0.8
0 1 2 3

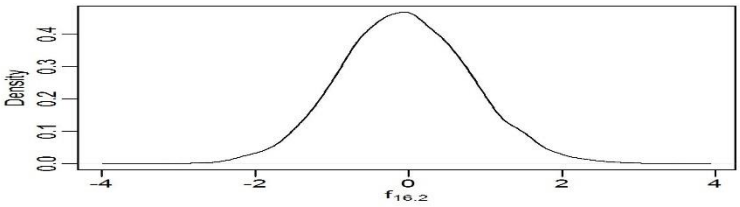

$f_{20.1}$
Iteractions
0 2000 4000 6000 8000 12000
-2 -1 0 1


Density
$f_{20.1}$
0.0 0.2 0.4 0.6 0.8 1.0
-3 -2 -1 0 1


$f_{14.2}$
Iteractions
0 2000 4000 6000 8000 12000
-4 -3 -2 -1 0 1 2 3


Density
$f_{14.2}$
0.0 0.1 0.2 0.3 0.4 0.5
-4 -2 0 2


$f_{16.2}$
Iteractions
0 2000 4000 6000 8000 12000
-2 0 2


Density
$f_{16.2}$
0.0 0.1 0.2 0.3 0.4
-4 -2 0 2 4

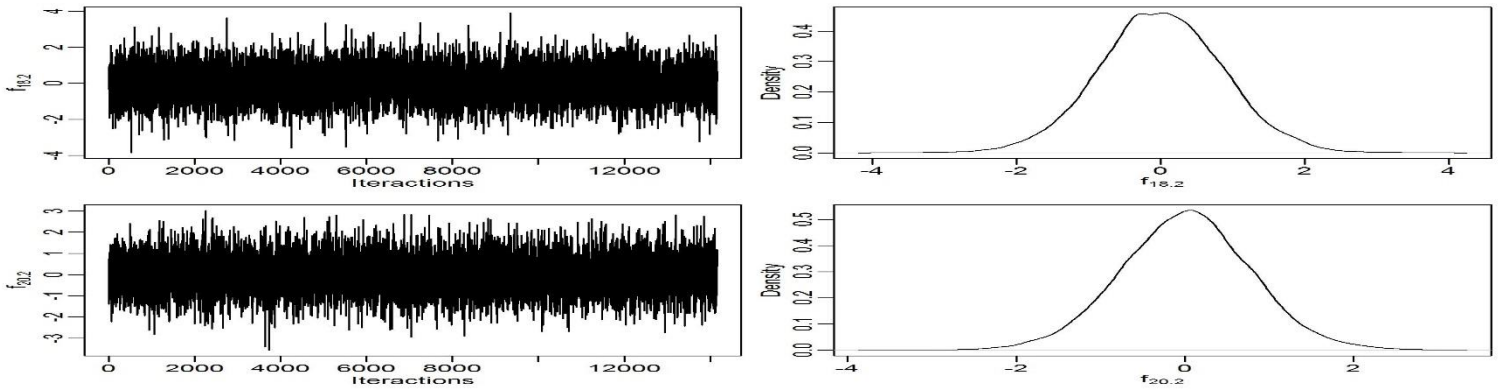
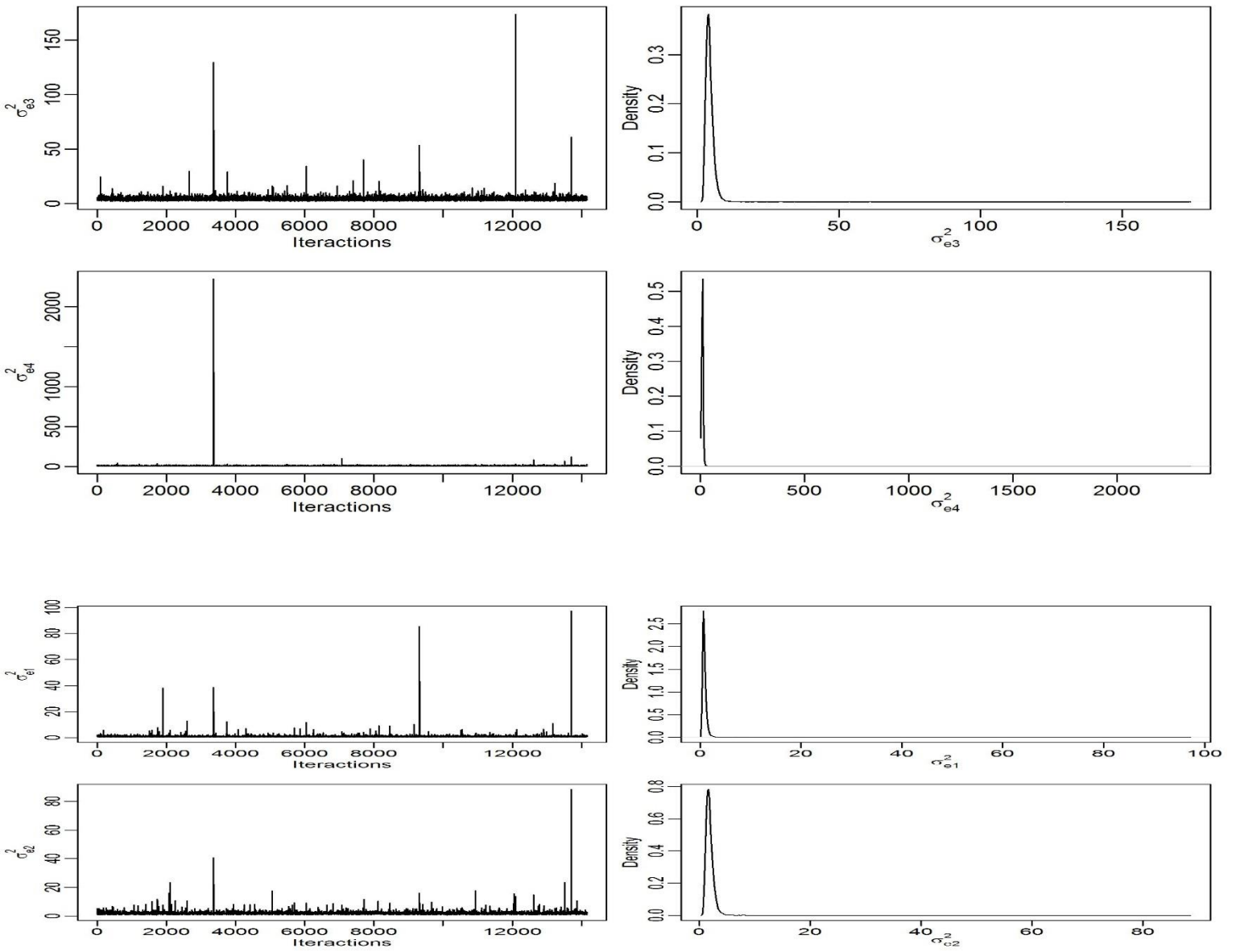
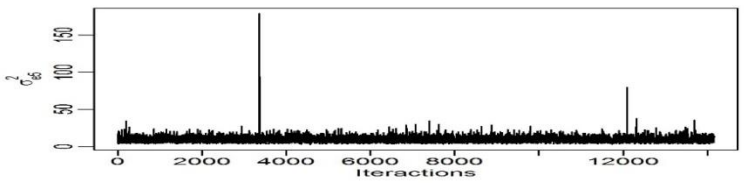
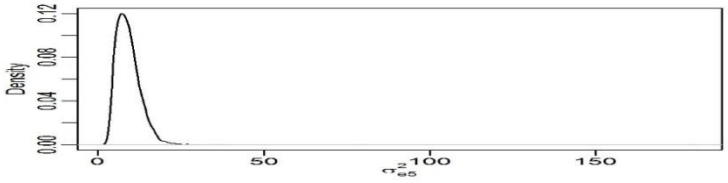


Figura 17-Traços das cadeias geradas pelo método MCMC e densidades a posteriori estimadas para as variâncias residuais.




σ^2_{es}
0
50
100
150
0
2000
4000
6000
8000
12000
Iteractions


Density
0.00
0.04
0.08
0.12
0
50
100
150
σ^2_{es}

APÊNDICE C-tabela resumo de inferências a posteriori dos efeitos de G e E

A seguir são apresentadas as tabelas com os resumos das inferências, a posteriori, para os efeitos principais de genótipos, para as coordenadas dos dois primeiros vetores relacionados aos dois primeiros eixos principais para os dados reais.

Tabela 6- Médias a posteriori (MP), regiões de credibilidade (I.C. 95%, LI: limite inferior, LS: limite superior para os dois primeiros vetores singulares ambientais ($\boldsymbol{\alpha}_{j1} - \boldsymbol{\alpha}_{j2}$) dos valores singulares ($\lambda_1 - \lambda_{10}$), e escores genotípicos ($\mathbf{f}_{i1} - \mathbf{f}_{i2}$) que não englobam a origem.

			HPD 95%					HPD 95%	
Par.	Média	sd	LI	LS	Par.	Média	sd	LI	LS
λ_2	1,113	0,175	0,800	1,480					
λ_5	0,581	0,126	0,346	0,840					
λ_6	0,442	0,107	0,230	0,643					
λ_7	0,338	0,108	0,108	0,548					
λ_8	0,183	0,110	0,000	0,370					
λ_9	0,085	0,075	0,000	0,236					
λ_{10}	0,040	0,046	0,000	0,133					
$\alpha_{2,1}$	0,295	0,073	0,137	0,424	$\alpha_{2,2}$	0,593	0,146	0,163	0,741
$\alpha_{5,1}$	0,282	0,072	0,140	0,420	$\alpha_{5,2}$	0,288	0,160	-0,119	0,516
$\alpha_{6,1}$	0,373	0,079	0,205	0,517	$\alpha_{6,2}$	-0,234	0,276	-0,675	0,404
$\alpha_{7,1}$	0,437	0,072	0,292	0,572	$\alpha_{7,2}$	-0,430	0,184	-0,666	0,055
$\alpha_{8,1}$	0,573	0,081	0,401	0,722	$\alpha_{8,2}$	0,160	0,252	-0,360	0,623
$\alpha_{9,1}$	0,204	0,063	0,081	0,329	$\alpha_{9,2}$	-0,207	0,160	-0,475	0,153
$\alpha_{10,1}$	0,323	0,081	0,165	0,481	$\alpha_{10,2}$	-0,255	0,234	-0,640	0,279
$f_{1,1}$	1,884	0,310	1,273	2,477	$f_{1,2}$	-0,400	0,995	-2,233	1,718
$f_{2,1}$	1,366	0,354	0,730	2,107	$f_{2,2}$	-0,651	0,805	-2,217	0,927
$f_{3,1}$	1,398	0,261	0,896	1,911	$f_{3,2}$	0,533	0,577	-0,560	1,704
$f_{5,1}$	1,659	0,312	1,069	2,289	$f_{5,2}$	-0,204	0,757	-1,689	1,282
$f_{6,1}$	1,251	0,255	0,752	1,754	$f_{6,2}$	0,990	0,517	-0,005	2,012
$f_{7,1}$	2,585	0,366	1,875	3,304	$f_{7,2}$	-0,418	0,851	-2,089	1,239
$f_{8,1}$	0,747	0,269	0,213	1,273	$f_{8,2}$	1,894	0,663	0,584	3,215
$f_{10,1}$	1,210	0,265	0,685	1,719	$f_{10,2}$	-0,906	0,640	-2,154	0,406
$f_{32,1}$	-1,460	0,509	-2,428	-0,429	$f_{32,2}$	0,641	0,871	-1,044	2,383
$f_{35,1}$	-2,285	0,517	-3,289	-1,291	$f_{35,2}$	0,754	0,841	-0,943	2,363
$f_{36,1}$	-1,349	0,494	-2,289	-0,359	$f_{36,2}$	0,148	0,744	-1,332	1,560
$f_{37,1}$	-1,417	0,475	-2,396	-0,541	$f_{37,2}$	0,165	0,733	-1,313	1,569
$f_{38,1}$	0,812	0,237	0,372	1,284	$f_{38,2}$	0,829	0,414	-0,008	1,632