



JÉSSICA GRACIELLE SILVA

**ABORDAGEM GEOMÉTRICA DOS PRINCIPAIS
DELINEAMENTOS EXPERIMENTAIS E DIAGRAMAS DE
HASSE**

LAVRAS - MG

2020

JÉSSICA GRACIELLE SILVA

**ABORDAGEM GEOMÉTRICA DOS PRINCIPAIS DELINEAMENTOS
EXPERIMENTAIS E DIAGRAMAS DE HASSE**

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

Prof. Dr. Lucas Monteiro Chaves

Orientador

LAVRAS - MG

2020

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

Silva, Jéssica Gracielle

Abordagem geométrica dos principais delineamentos
experimentais e diagramas de Hasse / Jéssica Gracielle Silva.
– Lavras : UFLA, 2020.

144 p. : il.

Dissertação(mestrado acadêmico)–Universidade Federal de
Lavras, 2020.

Orientador: Prof. Dr. Lucas Monteiro Chaves.

Bibliografia.

1. Análise de variância. 2. Teste de hipótese. 3. Delinea-
mentos ortogonais. I. Chaves, Lucas Monteiro. II. Título.

JÉSSICA GRACIELLE SILVA

**ABORDAGEM GEOMÉTRICA DOS PRINCIPAIS DELINEAMENTOS
EXPERIMENTAIS E DIAGRAMAS DE HASSE
AN GEOMETRIC APPROACH OF THE MAIN EXPERIMENTAL DESIGNS AND
HASSE DIAGRAMS**

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

APROVADA em 12 de Fevereiro de 2020.

Prof. Dr. Devanil Jaques de Souza UFLA
Prof. Dr. Renato Ribeiro de Lima UFLA
Prof. Dr. Eric Batista Ferreira UNIFAL

Prof. Dr. Lucas Monteiro Chaves
Orientador

**LAVRAS - MG
2020**

A Deus, que nos criou e foi tão dedicado nesta tarefa. Minha crença em sua existência foi meu sustento e me deu coragem para questionar a realidade e propor sempre um mundo novo de possibilidades. Sem Ele eu nada seria.
DEDICO.

AGRADECIMENTOS

Aos meus pais, Elizabeth e Paulo, por estarem sempre presentes em minha vida, pelo amor incondicional que me dedicaram. Tudo que sou devo a vocês.

À minha irmã Patrícia, pela amizade e pelo carinho que sempre me ofereceu.

Ao meu noivo e companheiro Thuler, pelo amor, carinho e incentivo durante a jornada.

Aos amigos que conheci durante o mestrado, por dividir tantos momentos bons.

Ao professor Dr. Lucas Monteiro Chaves, pela dedicação em me orientar e compartilhar comigo seus conhecimentos. Seus ensinamentos foram de extrema relevância na elaboração deste trabalho.

À Universidade Federal de Lavras, pela oportunidade de cursar o mestrado.

A todos os professores do Programa de Pós-graduação em Estatística e Experimentação Agropecuária, pelos ensinamentos transmitidos.

A todos os funcionários do Departamento de Estatística e do Departamento de Ciências Exatas, por serem sempre prestativos e amigos.

A todos que acreditaram em minha capacidade, vocês colaboraram para que a caminhada fosse mais leve.

O presente trabalho foi realizado com o apoio do Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq).

RESUMO

Seguindo exposição desenvolvida no livro "Design of comparative experiments" de Bailey (2008) a teoria dos delineamentos experimentais é apresentada, neste trabalho, com forte apelo geométrico. Conduziu-se, este trabalho, com o objetivo específico de abordar os tópicos apresentados por Bailey (2008) além de explorar ao máximo aspectos geométricos. Já, com o objetivo geral, pretende-se disponibilizar um texto amplo com valor didático. No presente trabalho, utiliza-se de metodologia de pesquisa bibliográfica alicerçada no livro de Bailey (2008). Com os resultados deste trabalho, busca-se alcançar os objetivos específico e geral citados. Para alcançar o objetivo específico, aspectos geométricos são fortemente levantados e também figuras geométricas são ilustradas em todo o texto. Para alcançar o objetivo geral, as demonstrações não apresentadas no livro de Bailey (2008) são realizadas em detalhes. Em decorrência da pouca difusão da abordagem geométrica dos delineamentos experimentais, este texto pretende ser, apesar da dificuldade de tal pretensão, uma referência didática em português a ser utilizada por estatísticos que planejam e analisam experimentos.

Palavras-chave: Análise de variância. Teste de hipótese. Delineamentos ortogonais.

ABSTRACT

Following an exposition developed in the book "Design of comparative experiments" by Bailey (2008), the theory of experimental designs is presented in the present study with a strong geometric appeal. Thus, this study aimed to explore to the maximum the geometric aspects approached by Bailey (2008). Generally, it is aimed to provide a broad text with didactic value. The present study uses a bibliographic research methodology based on Bailey's book (2008). The results of this study refer to achieving the specific and general objectives mentioned above. To achieve the specific objective, geometric aspects are strongly raised and also geometric figures are illustrated throughout the text. To achieve the general objective, the demonstrations not presented in Bailey's book (2008) are carried out in detail. Due to the little diffusion of the geometric approach of experimental designs, this text intends to be, despite the difficulty of such intention, a didactic reference in Portuguese to be used by statisticians who plan and analyze experiments.

Keywords: Analysis of variance. Hypothesis test. Orthogonal designs.

LISTA DE FIGURAS

Figura 2.1 – Função $T : \Omega \rightarrow \mathfrak{T}$	16
Figura 2.2 – Ângulo entre os vetores \mathbf{v} e \mathbf{w}	18
Figura 2.3 – Projeção de \mathbf{v} em \mathbf{w}	20
Figura 2.4 – Projeção de \mathbf{v} em W	21
Figura 2.5 – Espaço vetorial V	22
Figura 2.6 – Nuvem de pontos.	22
Figura 2.7 – Projeção de \mathbf{Y} em W	23
Figura 3.1 – Projeções de \mathbf{Y} em V_0 e W_T	32
Figura 3.2 – Projeção de \mathbf{Y} em V_T^\perp	33
Figura 3.3 – Decomposição de \mathbf{Y}	34
Figura 3.4 – Pitágoras.	35
Figura 3.5 – Pitágoras.	36
Figura 3.6 – Distribuição F	40
Figura 3.7 – Teste de hipótese para média geral.	41
Figura 3.8 – Teste de hipótese para tratamentos.	42
Figura 3.9 – Subespaços ortogonais para delineamento em blocos.	51
Figura 3.10 – Teste de hipótese para tratamentos.	56
Figura 3.11 – Teste de hipótese para blocos.	57
Figura 3.12 – Construção quadrado latino método cíclico.	59
Figura 3.13 – Construção do delineamento em linhas e colunas para o caso geral.	60
Figura 3.14 – Exemplo aleatorizado linhas e colunas.	60
Figura 3.15 – Subespaço V_T para o experimento fatorial.	76
Figura 3.16 – Teste de hipótese para o experimento fatorial.	82
Figura 3.17 – Interação entre variedades e métodos de cultivo.	83
Figura 3.18 – Subespaços para o princípio da ortogonalidade.	85
Figura 3.19 – Submodelos encaixados de $V_2 + V_2$	86
Figura 4.1 – Matriz \mathbf{X} como transformação linear.	89
Figura 4.2 – Subespaço \mathbb{R}^{t+b} restrito ao subespaço gerado por \mathbf{u}_0	94
Figura 6.1 – Fatores de tratamento F e G	113
Figura 6.2 – Fator de parcela T	114
Figura 6.3 – Croqui do experimento.	115

Figura 6.4 – Diagramas de Hasse para $F \prec G$ e $F \not\prec G$	118
Figura 6.5 – Diagramas de Hasse para fatores de tratamento e parcela.	118
Figura 6.6 – Diagramas de Hasse para o Exemplo 6.1.	119
Figura 6.7 – Graus de liberdade para o Exemplo 6.1.	119
Figura 6.8 – Coordenadas de V_F	120
Figura 6.9 – Subespaços ortogonais para os fatores F e G	121
Figura 6.10 – Subespaços ortogonais para o delineamento em blocos.	122
Figura 6.11 – Diagramas de Hasse para o delineamento inteiramente casualizado.	129
Figura 6.12 – Diagramas de Hasse para o delineamento em blocos casualizados.	130
Figura 6.13 – Diagramas de Hasse para o delineamento em linhas e colunas.	131
Figura 6.14 – Diagramas de Hasse para o experimento fatorial.	132
Figura 6.15 – Diagramas de Hasse para o Exemplo 6.6.1.	134
Figura 6.16 – Croqui do experimento para o exemplo 6.6.2.	135
Figura 6.17 – Diagramas de Hasse para o exemplo 6.6.2.	136

LISTA DE TABELAS

Tabela 3.1 – Vetores no experimento inteiramente casualizado.	29
Tabela 3.2 – Tabela de análise de variância para delineamento inteiramente casualizado.	37
Tabela 3.3 – Exemplo experimento em blocos.	48
Tabela 3.4 – Vetores no experimento em blocos.	49
Tabela 3.5 – Tabela de análise de variância para delineamento em blocos casualizados.	55
Tabela 3.6 – Exemplo experimento linhas e colunas.	58
Tabela 3.7 – Exemplo experimento em quadrado latino.	58
Tabela 3.8 – Vetores no experimento em linhas e colunas.	61
Tabela 3.9 – Tabela de análise de variância para delineamento em linhas e colunas.	67
Tabela 3.10 – Exemplo avaliação de vinhos.	70
Tabela 3.11 – Vetores no experimento fatorial.	72
Tabela 3.12 – Decomposição de V_T para experimento fatorial.	75
Tabela 3.13 – Quantidades associadas aos subespaços V	77
Tabela 3.14 – Tabela de análise de variância para experimento fatorial.	80
Tabela 5.1 – Tabela de análise de variância para delineamento inteiramente casualizado na presença de efeitos aleatórios.	102
Tabela 5.2 – Tabela de análise de variância para delineamento em blocos casualizados na presença de efeitos aleatórios.	105
Tabela 5.3 – Tabela de análise de variância para delineamento em linhas e colunas na presença de efeitos aleatórios.	109
Tabela 5.4 – Dados do experimento cordões de metal.	110
Tabela 5.5 – Tabela de análise de variância quando os cordões tem efeitos aleatórios.	111
Tabela 6.1 – Esquema de análise de variância para delineamento inteiramente casualizado por meio do diagrama de Hasse.	129
Tabela 6.2 – Esquema de análise de variância para delineamento em blocos casualizado por meio do diagrama de Hasse.	130
Tabela 6.3 – Esquema de análise de variância para delineamento em linhas e colunas por meio do diagrama de Hasse.	131
Tabela 6.4 – Esquema de análise de variância para experimento fatorial por meio do diagrama de Hasse.	132
Tabela 6.5 – Esquema de análise de variância para o Exemplo 6.6.1.	134

Tabela 6.6 – Esquema de análise de variância para o exemplo 6.6.2.	136
--	-----

SUMÁRIO

1	INTRODUÇÃO	12
2	UMA ABORDAGEM GEOMÉTRICA AOS DELINEAMENTOS EXPERIMENTAIS	15
2.1	Notação	15
2.2	Delineamentos experimentais	15
2.3	Vetores aleatórios	21
2.4	Análise de variância	26
3	O MÉTODO GEOMÉTRICO PARA ANÁLISE DE VARIÂNCIA EM DELINEAMENTOS EXPERIMENTAIS	28
3.1	Delineamento inteiramente casualizado	28
3.2	Delineamento em blocos casualizados	48
3.3	Delineamento em linhas e colunas	57
3.4	Experimento fatorial	69
3.5	Princípios de modelos de esperança	83
4	RELAÇÕES ENTRE A GEOMETRIA DOS MODELOS LINEARES E A GEOMETRIA DA ANÁLISE DE VARIÂNCIA	88
4.1	Modelos lineares na presença de efeitos de tratamentos	88
4.2	Modelos lineares na presença de efeitos de blocos	92
5	ANÁLISE DE VARIÂNCIA NA PRESENÇA DE EFEITOS ALEATÓRIOS	95
5.1	A métrica de Mahalanobis	95
5.2	Análise de variância na presença de covariância	100
5.2.1	Efeitos aleatórios no delineamento inteiramente casualizado	100
5.2.2	Efeitos aleatórios no delineamento em blocos casualizados	102
5.2.3	Efeitos aleatórios no delinemento em linhas e colunas	105
6	FUNDAMENTOS DA TEORIA DOS DIAGRAMAS DE HASSE APLICADOS À ANÁLISE DE VARIÂNCIA	112
6.1	Fatores, níveis, classes, equivalência e fatores especiais	112
6.2	Ínfimo e supremo	116
6.3	Diagramas de Hasse e graus de liberdade	117
6.4	Subespaços definidos por fatores, ortogonalidade e projeções	120
6.5	Delineamentos ortogonais	125

6.6	Análise de variância por meio do diagrama de Hasse	128
6.6.1	Exemplos de experimentos em parcelas subdivididas	132
7	METODOLOGIA	137
8	RESULTADOS	138
9	CONCLUSÃO	140
10	REFERÊNCIAS	141
	APENDICE A – DEFINIÇÕES E TEOREMAS RELEVANTES	142

1 INTRODUÇÃO

No presente trabalho, apresenta-se uma abordagem geométrica da teoria dos delineamentos experimentais. Tal abordagem é descrita no livro “Design of comparative experiments” de autoria Bailey (2008).

Conduziu-se, este trabalho, com o objetivo específico de abordar os tópicos citados por Bailey (2008), além de explorar o máximo dos aspectos geométricos. Já, este trabalho, tem como objetivo geral disponibilizar um texto amplo com valor didático. Para tanto, utilizaram-se de figuras geométricas para visualização de conceitos essenciais no desenvolver do trabalho. Utilizaram-se também conceitos geométricos tais como: espaços vetoriais, projeções ortogonais, funções definidas por fatores e diagramas de Hasse.

Em Bailey (2008), as demonstrações de alguns teoremas não são apresentadas. Com o propósito de disponibilizar um texto amplo, no presente trabalho, expõem-se as demonstrações dos teoremas do livro que foram utilizados e que não possuem suas respectivas demonstrações. Já, as demonstrações dos teoremas que se encontram em Bailey (2008) não foram apresentadas neste trabalho.

Pretende-se trazer uma abordagem geométrica e explicitar a álgebra associada à análise de variância para diferentes delineamentos, enfatizando, sempre que possível, seu aspecto intuitivo. O uso da geometria pode deixar explícitos conceitos estatísticos, bem como permitir a compreensão da estatística em sua totalidade.

A abordagem geométrica, apesar de ser clássica e utilizada pelos pioneiros da área não é muito difundida no ensino acadêmico. Em geral, uma abordagem bastante algébrica é utilizada no ensino de análise de experimentos. O método geométrico possui a vantagem de permitir melhor compreensão a quem tem familiaridade com tais conceitos e justifica um texto que torne essa abordagem clara e didaticamente acessível aos estudantes.

Nesse sentido, os principais delineamentos experimentais são abordados neste texto com enfoque geométrico. Esses delineamentos são: delineamento inteiramente casualizado, delineamento em blocos casualizados e delineamento em linhas e colunas. Com dados provenientes de um experimento conduzido em algum desses delineamentos é possível utilizar a técnica estatística conhecida como análise de variância.

A análise de variância é uma das técnicas estatísticas mais amplamente utilizadas. Uma ideia básica da análise de variância, a do particionamento da variabilidade, é fundamental para a estatística experimental. A análise de variância desmente seu próprio nome, no sentido que não está preocupada em analisar

variâncias, mas sim, a variabilidade nas médias (CASELLA; BERGER, 2010, p. 465).

Esse particionamento da variabilidade é apresentado, neste texto, como a decomposição do espaço de dados em soma direta de subespaços e, conseqüentemente, do vetor de respostas em vetores correspondentes a esses subespaços.

Primeiramente, conceitos usuais sobre delineamentos experimentais são apresentados, bem como alguns teoremas que envolvem conceitos geométricos que são essenciais no desenvolver do trabalho. Em seguida, são expostas características de vetores aleatórios e também definições para se utilizar a técnica de análise de variância, enfatizando sempre conceitos geométricos.

Posteriormente, os delineamentos experimentais são apresentados com enfoque geométrico. Para tanto, o espaço vetorial de resposta é decomposto em subespaços e são obtidas as projeções do vetor de respostas nesses subespaços. A partir da norma ao quadrado desses vetores de projeção, é possível obter as somas de quadrados da análise de variância. Cada caso particular de delineamento resulta em decomposições diferentes do espaço vetorial de respostas. Nessa abordagem são apresentados os diferentes tipos de delineamentos experimentais.

No delineamento inteiramente casualizado, os tratamentos são alocados aleatoriamente às parcelas experimentais. Esse tipo de delineamento exige área experimental homogênea. Geralmente, experimentos planejados no delineamento inteiramente casualizado possuem grande número de unidades amostrais.

No delineamento em blocos casualizados, os tratamentos são alocados aleatoriamente às parcelas em cada bloco. Nesse tipo de delineamento, a área experimental não é totalmente homogênea. Desse modo, é necessário agrupar as parcelas em blocos que sejam mais homogêneos. Cada bloco representa um conjunto de parcelas homogêneas do experimento. Esse agrupamento dos blocos é denominado controle local. Se cada bloco recebe todos os tratamentos, o experimento é denominado em blocos completos.

Já, no delineamento em linhas e colunas, os tratamentos são alocados aleatoriamente às linhas e as colunas. Nesse tipo de delineamento, a área experimental possui duas fontes de variação não homogêneas que são ortogonias. Desse modo, o experimento necessita de controle local duplo.

Em seguida, é apresentado o experimento fatorial, nele os tratamentos são todas as combinações dos níveis de fatores em estudo. O experimento fatorial não é um delineamento particular e pode ser planejado nos três delineamentos mencionados anteriormente.

Após apresentar os tipos usuais de delineamentos uma comparação é exposta entre a abordagem apresentada, neste texto, e a abordagem geralmente utilizada nos livros didáticos.

Posteriormente, a análise de variância é expressa na presença de efeitos aleatórios. Nesse momento, a decomposição do espaço de dados é mostrada em relação à autoespaços e consequentemente a autovetores associados a autovalores distintos.

Em seguida, é apresentado o capítulo referente aos diagramas de Hasse. Nesse é formalizado o conceito de fatores, que são funções que definem partições. Essas partições são definidas no conjunto de tratamentos ou no conjunto de parcelas. Determinados os fatores de um delineamento experimental, os diagramas de Hasse podem ser definidos e toda a análise de variância é obtida.

Por fim, são apresentados os capítulos referentes à metodologia, resultados e conclusão. No capítulo metodológico, relata-se como o trabalho foi realizado e os passos seguidos, durante a escrita. No capítulo dos resultados, apresentam-se, como os objetivos da pesquisa foram alcançados. E o capítulo de conclusão traz uma síntese de tudo que foi realizado.

2 UMA ABORDAGEM GEOMÉTRICA AOS DELINEAMENTOS EXPERIMENTAIS

2.1 Notação

No presente trabalho, vetores de variáveis aleatórias bem como vetores de realizações dessas variáveis aleatórias são ambos indicados por \mathbf{Y} .

Vetores escalares são denotados por letras minúsculas em negrito \mathbf{v} e \mathbf{w} . O transposto desses vetores por \mathbf{v}^\top e \mathbf{w}^\top .

Matrizes são indicadas por letras maiúsculas em negrito \mathbf{X} e \mathbf{I} . A tranposta dessas matrizes é indicada por \mathbf{X}^\top e \mathbf{I}^\top e a inversa por \mathbf{X}^{-1} e \mathbf{I}^{-1} .

Conjuntos são denotados por letras maiúsculas V e W . O complemento ortogonal desses conjuntos por V^\perp e W^\perp .

Espaços vetoriais ¹ e o complemento ortogonal desses espaços vetoriais são indicados da mesma maneira que conjuntos.

Funções e fatores são denotados por letras maiúsculas T e F .

Por fim, a dimensão de um espaço vetorial V é indicada por $\dim(V)$.

Antes de iniciar, é necessário enfatizar que a maioria das definições e teoremas presentes, neste trabalho, encontram-se no livro de Bailey (2008). O livro citado é a principal referência teórica deste trabalho.

2.2 Delineamentos experimentais

No que se refere a delineamentos experimentais, alguns conceitos são de extrema relevância, os quais são apresentados nesta seção, conforme constam em Bailey (2008).

Uma unidade experimental é a menor unidade na qual se aloca um tratamento. Uma unidade observacional é a menor unidade na qual a resposta pode ser medida. Uma parcela é quando a unidade experimental é igual à unidade observacional ou quando a unidade experimental possui várias unidades observacionais. Um tratamento é tudo que pode ser alocado em uma parcela com a finalidade de ser avaliado ou estudado.

Neste trabalho, parcelas são indicadas por letras gregas minúsculas $\alpha, \beta, \gamma, \omega$. Tratamentos são indicados por letras latinas minúsculas i, j, k . O conjunto de todas as parcelas de um experimento é denotado por Ω . O número de parcelas por N . O conjunto de todos os tratamentos é denotado por \mathfrak{T} . O número total de tratamentos por t . O número de vezes que cada

¹ Veja Definição .0.1 do Apêndice deste trabalho.

tratamento i é aplicado às parcelas é indicado por r_i . Assim, se o tratamento i foi reaplicado r_i vezes, temos que:

$$\sum_{i=1}^t r_i = |\Omega| = N,$$

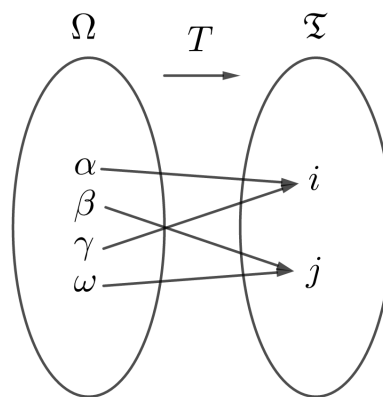
ou seja, o somatório do número de repetições dos tratamentos resulta no número de parcelas ².

A estrutura de parcela, é definida como formas significativas de dividir o conjunto Ω . A estrutura de parcela dita como será o delineamento experimental, que pode ser: inteiramente casualizado (não existe estrutura nas parcelas, ou seja, as parcelas não são agrupadas em blocos); em blocos casualizados (as parcelas são divididas em blocos homogêneos); dentre outros. A estrutura de tratamento significa formas significativas de dividir o conjunto \mathfrak{T} .

O delineamento experimental, ou *design*, é a maneira de alocar os tratamentos às parcelas. Matematicamente, o design é uma função de Ω em \mathfrak{T} , isso porque cada parcela pode receber apenas um tratamento. Essa função é denotada $T : \Omega \rightarrow \mathfrak{T}$. Desse modo, $T(\omega)$ é o tratamento alocado na parcela ω e $Y(\omega)$ é a resposta.

Exemplificando na Figura 2.1, o delineamento experimental como função T de $\Omega = \{\alpha, \beta, \gamma, \omega\}$ em $\mathfrak{T} = \{i, j\}$. Note que, $N = 4$, $t = 2$, e conseqüentemente, $r_i = 2$ e $r_j = 2$.

Figura 2.1 – Função $T : \Omega \rightarrow \mathfrak{T}$.



Fonte: Da autora (2019).

Como já mencionado anteriormente, o conjunto das parcelas é denotado por Ω . Esse conjunto está associado a um espaço vetorial N-dimensional, que consiste de uma sequência finita de números reais, com cada local na sequência associado a uma parcela. Como $|\Omega| = N$, precisamente, esse espaço vetorial é $\mathbb{R}^{|\Omega|} = \mathbb{R}^N$ e assim como em Bailey (2008) será denotado

² $|\Omega|$ é a cardinalidade do conjunto, ou seja, a quantidade de elementos de Ω .

por V . Os subespaços ³ de V são de suma importância para obter a análise de variância, como será visto mais adiante neste trabalho.

A base de qualquer espaço vetorial V é caracterizada por um conjunto de vetores que geram esse espaço. Todo vetor \mathbf{v} em V é uma combinação linear ⁴ única dos vetores da base.

Para exemplificar, o conjunto de vetores $\{(1, 0), (0, 1)\}$ formam uma base para o espaço vetorial \mathbb{R}^2 , uma vez que, esse conjunto possui dois vetores, $\dim(\mathbb{R}^2) = 2$.

A seguir, algumas definições e teoremas necessários no desenvolver deste trabalho.

Definição 2.2.1. *O produto escalar ou produto interno entre dois vetores \mathbf{v} e \mathbf{w} em V é definido por:*

$$\mathbf{v} \cdot \mathbf{w} = \sum_{\omega \in \Omega} v_{\omega} w_{\omega} = v_1 w_1 + v_2 w_2 + \dots + v_N w_N = \mathbf{v}^T \mathbf{w}.$$

Definição 2.2.2. *A norma ao quadrado de um vetor \mathbf{v} é definida por:*

$$\|\mathbf{v}\|^2 = \mathbf{v} \cdot \mathbf{v} = \sum_{\omega \in \Omega} v_{\omega} v_{\omega} = \sum_{\omega \in \Omega} v_{\omega}^2 = v_1^2 + v_2^2 + \dots + v_N^2 = \mathbf{v}^T \mathbf{v}.$$

Definição 2.2.3. *Um vetor \mathbf{v} é unitário se $\|\mathbf{v}\| = \sqrt{\mathbf{v} \cdot \mathbf{v}} = 1$.*

Teorema 2.2.1. *O produto interno entre dois vetores \mathbf{v} e \mathbf{w} satisfaz:*

$$\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\| \|\mathbf{w}\| \cos(\theta),$$

em que θ é o ângulo formado entre os vetores \mathbf{v} e \mathbf{w} .

Demonstração. Da lei dos cossenos sabemos que: $\|\mathbf{v} - \mathbf{w}\|^2 = \|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\| \|\mathbf{w}\| \cos(\theta)$, conforme ilustração na Figura 2.2.

³ Veja Definição .0.2 do Apêndice deste trabalho.

⁴ Veja Definição .0.3 do Apêndice deste trabalho.

Desenvolvendo o primeiro membro da equação temos:

$$\begin{aligned}
 \|\mathbf{v} - \mathbf{w}\|^2 &= (\mathbf{v} - \mathbf{w}) \cdot (\mathbf{v} - \mathbf{w}) \\
 &= (v_1 - w_1)^2 + (v_2 - w_2)^2 + \dots + (v_N - w_N)^2 \\
 &= v_1^2 - 2v_1w_1 + w_1^2 + v_2^2 - 2v_2w_2 + w_2^2 + \dots + v_N^2 - 2v_Nw_N + w_N^2 \\
 &= v_1^2 + v_2^2 + \dots + v_N^2 + w_1^2 + w_2^2 + \dots + w_N^2 - 2v_1w_1 - 2v_2w_2 - \dots - 2v_Nw_N \\
 &= (v_1^2 + v_2^2 + \dots + v_N^2 + w_1^2 + w_2^2 + \dots + w_N^2) - 2(v_1w_1 + v_2w_2 + \dots + v_Nw_N) \\
 &= (v_1^2 + v_2^2 + \dots + v_N^2) + (w_1^2 + w_2^2 + \dots + w_N^2) - 2(\mathbf{v} \cdot \mathbf{w}) \\
 &= \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2(\mathbf{v} \cdot \mathbf{w}).
 \end{aligned}$$

Desenvolvendo o segundo membro da equação temos:

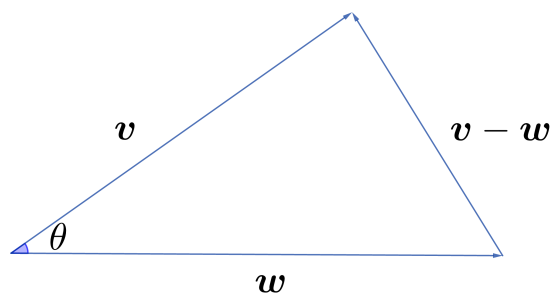
$$\|\mathbf{v}\|^2 + \|\mathbf{w}\|^2 - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta) = \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta).$$

Igualando ambos os membros da equação obtemos:

$$\begin{aligned}
 \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2(\mathbf{v} \cdot \mathbf{w}) &= \mathbf{v} \cdot \mathbf{v} + \mathbf{w} \cdot \mathbf{w} - 2\|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta) \\
 -2(\mathbf{v} \cdot \mathbf{w}) &= -2\|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta) \\
 \mathbf{v} \cdot \mathbf{w} &= \|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta).
 \end{aligned}$$

□

Figura 2.2 – Ângulo entre os vetores \mathbf{v} e \mathbf{w} .



Fonte: Da autora (2019).

Definição 2.2.4. Dois vetores \mathbf{v} e \mathbf{w} são ortogonais se $\mathbf{v} \cdot \mathbf{w} = 0$. Escrevemos $\mathbf{v} \perp \mathbf{w}$.

Note que se $\mathbf{v} \cdot \mathbf{w} = 0$ para dois vetores $\mathbf{v} \neq \mathbf{0}$ e $\mathbf{w} \neq \mathbf{0}$, temos que $\|\mathbf{v}\| \neq 0$ e $\|\mathbf{w}\| \neq 0$. Como o produto interno satisfaz $\|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta)$, nesse caso temos que $\cos(\theta) = 0$, logo $\theta = 90^\circ$ é o ângulo formado entre \mathbf{v} e \mathbf{w} .

Teorema 2.2.2. *A projeção ortogonal de um vetor \mathbf{v} na direção de um vetor \mathbf{w} é dada por:*

$$P_{\mathbf{w}}\mathbf{v} = \left(\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \right) \mathbf{w}.$$

Demonstração. Da trigonometria no triângulo retângulo sabemos que $\cos(\theta) = \frac{\|P_{\mathbf{w}}\mathbf{v}\|}{\|\mathbf{v}\|}$, Figura 2.3. Sendo assim:

$$\|P_{\mathbf{w}}\mathbf{v}\| = \|\mathbf{v}\|\cos(\theta).$$

Como o produto interno satisfaz $\mathbf{v} \cdot \mathbf{w} = \|\mathbf{v}\|\|\mathbf{w}\|\cos(\theta)$ temos que:

$$\begin{aligned} \|P_{\mathbf{w}}\mathbf{v}\| &= \|\mathbf{v}\| \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{v}\|\|\mathbf{w}\|} \\ &= \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|}. \end{aligned}$$

Multiplicando $\|P_{\mathbf{w}}\mathbf{v}\|$ pelo vetor unitário ⁵ $\frac{\mathbf{w}}{\|\mathbf{w}\|}$, obtemos:

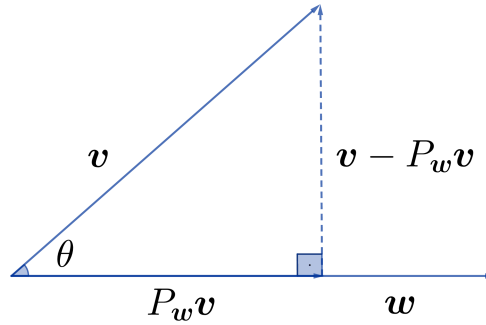
$$\begin{aligned} \|P_{\mathbf{w}}\mathbf{v}\| \frac{\mathbf{w}}{\|\mathbf{w}\|} &= \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|} \frac{\mathbf{w}}{\|\mathbf{w}\|} \\ &= \frac{\mathbf{v} \cdot \mathbf{w}}{\|\mathbf{w}\|^2} \mathbf{w} \\ &= \left(\frac{\mathbf{v} \cdot \mathbf{w}}{\mathbf{w} \cdot \mathbf{w}} \right) \mathbf{w} \\ &= P_{\mathbf{w}}\mathbf{v}. \end{aligned}$$

□

Note que, por subtração de vetores, é possível obter $\mathbf{v} - P_{\mathbf{w}}\mathbf{v}$. Note também que $P_{\mathbf{w}}\mathbf{v}$ é um múltiplo escalar do vetor \mathbf{w} , assim como mostra a Figura 2.3.

Definição 2.2.5. *Se W é um subespaço de V , o complemento ortogonal de W é o conjunto $W^\perp = \{\mathbf{v} \in V : \mathbf{v} \perp \mathbf{w}, \forall \mathbf{w} \in W\}$.*

⁵ Veja Definição .0.4 do Apêndice deste trabalho.

Figura 2.3 – Projeção de \mathbf{v} em \mathbf{w} .

Fonte: Da autora (2019).

Teorema 2.2.3. *Seja W um subespaço d -dimensional de V . Se $\{\mathbf{u}_1, \dots, \mathbf{u}_d\}$ é uma base ortogonal⁶ de W , então a projeção ortogonal do vetor \mathbf{v} no subespaço W é dada por:*

$$P_W \mathbf{v} = \left(\frac{\mathbf{v} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \dots + \left(\frac{\mathbf{v} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d} \right) \mathbf{u}_d.$$

Demonstração. Reescrevendo $P_W \mathbf{v}$ como $\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i$. Usando o fato de que \mathbf{v} é um vetor de V e pode ser escrito como soma de um vetor em W e outro em W^\perp , ou seja, $\mathbf{v} = P_W \mathbf{v} + \mathbf{h}$. Basta mostrar que $P_W \mathbf{v} \cdot (\mathbf{v} - P_W \mathbf{v}) = 0$, pois se isso acontece esses vetores são perpendiculares e P_W é o projetor. Vejamos:

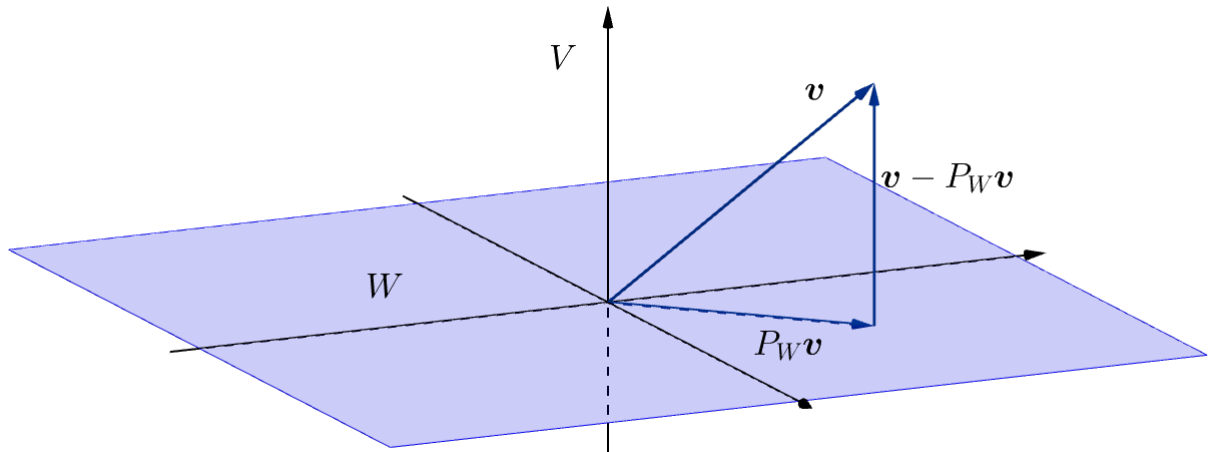
$$\begin{aligned} P_W \mathbf{v} \cdot (\mathbf{v} - P_W \mathbf{v}) &= P_W \mathbf{v} \cdot \mathbf{v} - P_W \mathbf{v} \cdot P_W \mathbf{v} \\ &= P_W \mathbf{v} \cdot \mathbf{v} - \|P_W \mathbf{v}\|^2 \\ &= \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \right) \cdot (P_W \mathbf{v} + \mathbf{h}) - \|P_W \mathbf{v}\|^2 \\ &= \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \right) \cdot \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i + \mathbf{h} \right) - \|P_W \mathbf{v}\|^2 \\ &= \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \right) \cdot \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \right) + \left(\sum_{i=1}^d \left(\frac{\mathbf{v} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \right) \cdot \mathbf{h} - \|P_W \mathbf{v}\|^2 \\ &= P_W \mathbf{v} \cdot P_W \mathbf{v} + 0 - \|P_W \mathbf{v}\|^2 \\ &= \|P_W \mathbf{v}\|^2 - \|P_W \mathbf{v}\|^2 \\ &= 0 \end{aligned}$$

Portanto, $P_W \mathbf{v}$ é o projetor. □

⁶ Veja Definição .0.5 do Apêndice deste trabalho.

Novamente, por subtração de vetores, é possível obter $\mathbf{v} - P_W \mathbf{v}$, assim como mostra a Figura 2.4.

Figura 2.4 – Projeção de \mathbf{v} em W .



Fonte: Da autora (2019).

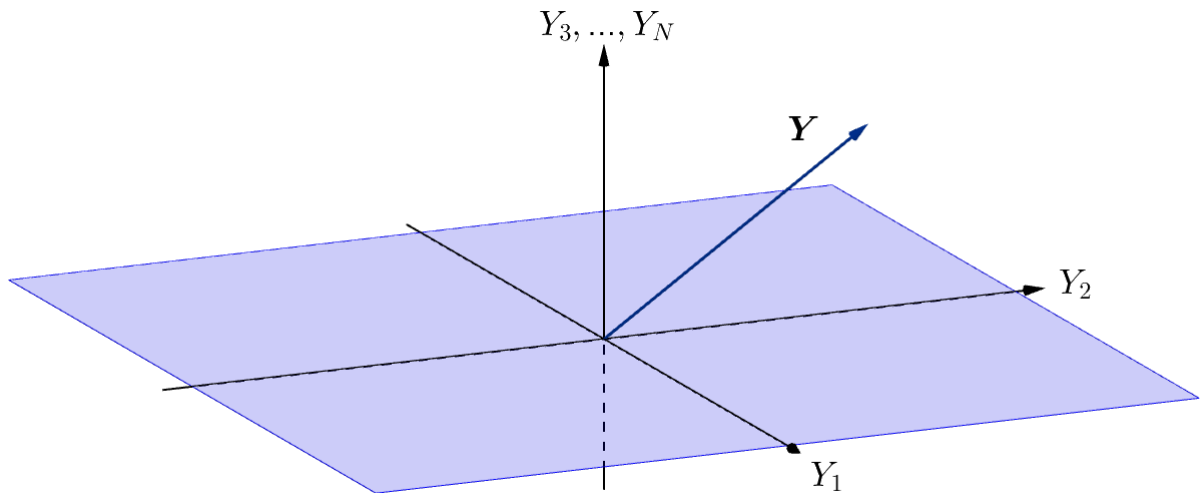
As definições e teoremas apresentados podem ser estendidos a vetores aleatórios. Com isso, é possível utilizar a técnica estatística de análise de variância, utilizando conceitos geométricos.

2.3 Vetores aleatórios

Para um experimento, tem-se um vetor de variáveis aleatórias. Cada realização dessas variáveis aleatórias resulta em um vetor de respostas. Esse vetor de respostas é um vetor do espaço vetorial V , veja Figura 2.5. No presente trabalho, vetor de variáveis aleatórias e vetor de respostas são ambos indicados por $\mathbf{Y}^T = (Y_1, Y_2, \dots, Y_N)$.

Variáveis aleatórias possuem função densidade de probabilidade. Para o caso multivariado \mathbf{Y} , é função de N variáveis. Uma função densidade de probabilidade essencial em estatística é a normal, uma vez que, a maioria dos dados provenientes de populações segue distribuição normal. Para o caso em que \mathbf{Y} possui densidade de probabilidade normal multivariada a função é dada por:

$$f(\mathbf{Y}) = \frac{1}{(\sqrt{2\pi})^N |\boldsymbol{\Sigma}|^{\frac{1}{2}}} e^{-\frac{1}{2}(\mathbf{Y}-\boldsymbol{\tau})^T(\boldsymbol{\Sigma})^{-1}(\mathbf{Y}-\boldsymbol{\tau})},$$

Figura 2.5 – Espaço vetorial V .

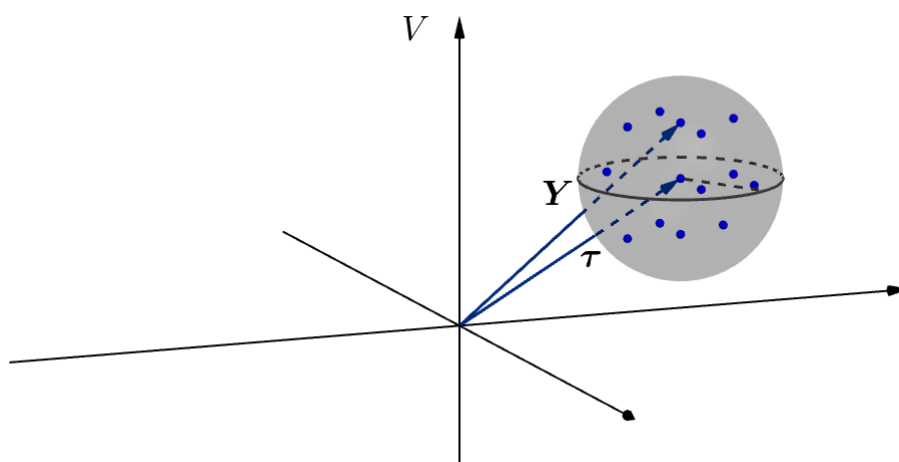
Fonte: Da autora (2019).

em que $\boldsymbol{\tau}$ é o vetor de médias ($N \times 1$), $\boldsymbol{\Sigma}$ é uma matriz simétrica de variâncias e covariâncias ($N \times N$) de \mathbf{Y} e $|\boldsymbol{\Sigma}|$ é o determinante da matriz $\boldsymbol{\Sigma}$.

Um caso particular é considerado, quando as realizações das variáveis aleatórias de \mathbf{Y} são independentes. Nesse caso, a matriz de variâncias e covariâncias $\boldsymbol{\Sigma}$ é dada por $\sigma^2 \mathbf{I}$, em que σ^2 é a variância e \mathbf{I} é a matriz identidade de ordem N , o que significa que a correlação entre duas variáveis aleatórias distintas é nula.

Assuma que $\mathbf{Y} \sim N(\boldsymbol{\tau}, \sigma^2 \mathbf{I})$, nesse caso as realizações de \mathbf{Y} tendem a ser centradas em um vetor de médias $\boldsymbol{\tau}$ e podem ser representadas por uma nuvem de pontos de formato esférico, como mostra a Figura 2.6.

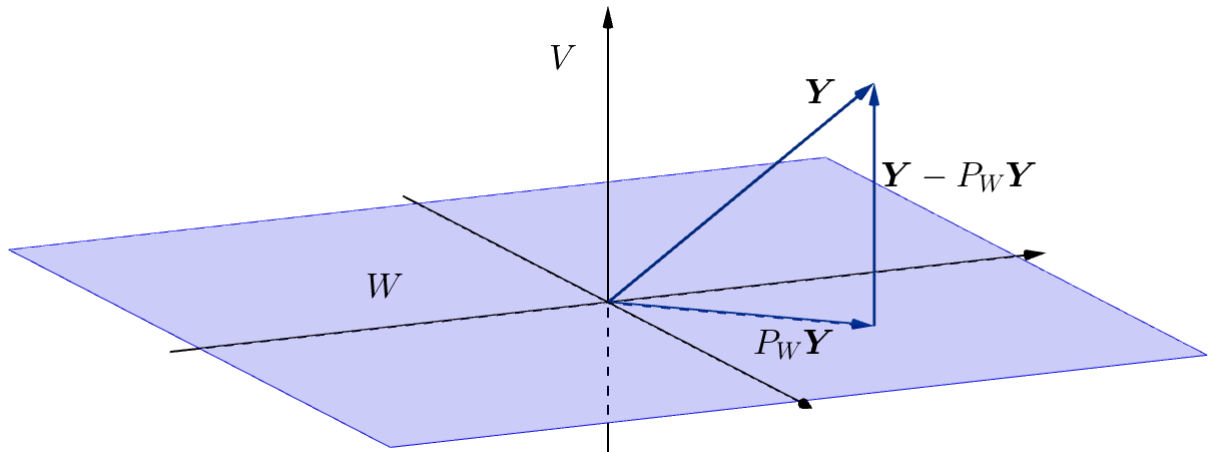
Figura 2.6 – Nuvem de pontos.



Fonte: Da autora (2019).

Lembre-se da projeção ortogonal de vetores, essas projeções são estendidas a vetores aleatórios. Sendo \mathbf{Y} um vetor aleatório de um espaço vetorial V , a projeção ortogonal de \mathbf{Y} em qualquer subespaço W de V também é um vetor aleatório que depende de \mathbf{Y} (Figura 2.7).

Figura 2.7 – Projeção de \mathbf{Y} em W .



Fonte: Da autora (2019).

Teorema 2.3.1. *Suponha que $E(\mathbf{Y}) = \boldsymbol{\tau}$ e que $\text{cov}(\mathbf{Y}) = \mathbf{I}\sigma^2$. Seja W um subespaço d -dimensional de V , Então*

$$i) E(P_W \mathbf{Y}) = P_W(E(\mathbf{Y})) = P_W \boldsymbol{\tau};$$

$$ii) E(\|P_W \mathbf{Y}\|^2) = \|P_W \boldsymbol{\tau}\|^2 + d\sigma^2.$$

Demonstração.

i) Seja $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d\}$ uma base ortogonal de W . Expressando $P_W \mathbf{Y}$ com relação a essa base e, em seguida, calculando a esperança desse projetor obtemos:

$$P_W \mathbf{Y} = \left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d} \right) \mathbf{u}_d,$$

$$\begin{aligned}
E(P_W \mathbf{Y}) &= E\left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d\right) \\
&= E\left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1\right) + E\left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2\right) + \dots + E\left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d\right) \\
&= \left(\frac{E(\mathbf{Y}) \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{E(\mathbf{Y}) \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 + \dots + \left(\frac{E(\mathbf{Y}) \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d \\
&= \left(\frac{\boldsymbol{\tau} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{\boldsymbol{\tau} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 + \dots + \left(\frac{\boldsymbol{\tau} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d \\
&= P_W \boldsymbol{\tau}.
\end{aligned}$$

Portanto, $E(P_W \mathbf{Y}) = P_W \boldsymbol{\tau}$.

ii) Seja $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_d\}$ uma base ortonormal⁷ de W . Expressando $P_W \mathbf{Y}$ com relação a essa base e calculando a norma ao quadrado obtemos:

$$\begin{aligned}
\|P_W \mathbf{Y}\|^2 &= P_W \mathbf{Y} \cdot P_W \mathbf{Y} \\
&= \left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d\right) \cdot \\
&\quad \left(\left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1}\right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2}\right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_d}{\mathbf{u}_d \cdot \mathbf{u}_d}\right) \mathbf{u}_d\right) \\
&= ((\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_d) \mathbf{u}_d) \cdot \\
&\quad ((\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_d) \mathbf{u}_d) \\
&= (\mathbf{Y} \cdot \mathbf{u}_1)^2 + (\mathbf{Y} \cdot \mathbf{u}_2)^2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_d)^2.
\end{aligned}$$

Calculando agora a esperança desse projetor temos:

$$\begin{aligned}
E(\|P_W \mathbf{Y}\|^2) &= E((\mathbf{Y} \cdot \mathbf{u}_1)^2 + (\mathbf{Y} \cdot \mathbf{u}_2)^2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_d)^2) \\
&= E((\mathbf{Y} \cdot \mathbf{u}_1)^2) + E((\mathbf{Y} \cdot \mathbf{u}_2)^2) + \dots + E((\mathbf{Y} \cdot \mathbf{u}_d)^2) \\
&= \sum_{i=1}^d E((\mathbf{Y} \cdot \mathbf{u}_i)^2).
\end{aligned}$$

Como $\text{var}(Y) = E(Y^2) - (E(Y))^2$ temos que:

$$\begin{aligned}
\text{var}(\mathbf{Y} \cdot \mathbf{u}_i) &= E((\mathbf{Y} \cdot \mathbf{u}_i)^2) - (E(\mathbf{Y} \cdot \mathbf{u}_i))^2 \implies \\
E((\mathbf{Y} \cdot \mathbf{u}_i)^2) &= \text{var}(\mathbf{Y} \cdot \mathbf{u}_i) + (E(\mathbf{Y} \cdot \mathbf{u}_i))^2.
\end{aligned}$$

⁷ Veja Definição .0.6 do Apêndice deste trabalho.

Logo,

$$\begin{aligned}
E(\|P_W \mathbf{Y}\|^2) &= \sum_{i=1}^d (\text{var}(\mathbf{Y} \cdot \mathbf{u}_i) + (E(\mathbf{Y} \cdot \mathbf{u}_i))^2) \\
&= \sum_{i=1}^d \text{var}(\mathbf{u}_i \cdot \mathbf{Y}) + \sum_{i=1}^d (E(\mathbf{u}_i \cdot \mathbf{Y}))(E(\mathbf{u}_i \cdot \mathbf{Y})) \\
&= \sum_{i=1}^d \|\mathbf{u}_i\|^2 \sigma^2 + \sum_{i=1}^d (\mathbf{u}_i \cdot \boldsymbol{\tau})(\mathbf{u}_i \cdot \boldsymbol{\tau}) \\
&= \sum_{i=1}^d \mathbf{u}_i \cdot \mathbf{u}_i \sigma^2 + \sum_{i=1}^d (\mathbf{u}_i \cdot \boldsymbol{\tau})^2 \\
&= \sum_{i=1}^d 1 \sigma^2 + \sum_{i=1}^d (\mathbf{u}_i \cdot \boldsymbol{\tau})^2 \\
&= d\sigma^2 + (\mathbf{u}_1 \cdot \boldsymbol{\tau})^2 + (\mathbf{u}_2 \cdot \boldsymbol{\tau})^2 + \dots + (\mathbf{u}_d \cdot \boldsymbol{\tau})^2 \\
&= d\sigma^2 + \|P_W \boldsymbol{\tau}\|^2.
\end{aligned}$$

Portanto, $E(\|P_W \mathbf{Y}\|^2) = \|P_W \boldsymbol{\tau}\|^2 + d\sigma^2$.

A passagem $\text{var}(\mathbf{u}_i \cdot \mathbf{Y}) = \|\mathbf{u}_i\|^2 \sigma^2$ será vista mais adiante, no Teorema 3.1.3, quando se inserir o conceito de estimadores.

□

Como se pretende obter um texto didático, outra demonstração para esse teorema é apresentada a seguir.

Considere uma base de V dada por $\{\mathbf{u}_1, \dots, \mathbf{u}_d, \mathbf{w}_{d+1}, \dots, \mathbf{w}_N\}$ ortonormal. Note que essa base é uma extensão da base de W . Como sabemos, $\mathbf{Y}^\top = (Y_1, \dots, Y_N)$ e para qualquer vetor dessa base, $\mathbf{Y} \cdot \mathbf{u}_i = Y_i$. Sendo assim, temos que $E((\mathbf{Y} \cdot \mathbf{u}_i)^2) = E(Y_i^2)$. Calculando a norma ao quadrado do projetor e, em seguida, a esperança obtemos:

$$\begin{aligned}
\|P_W \mathbf{Y}\|^2 &= (\mathbf{Y} \cdot \mathbf{u}_1)^2 + (\mathbf{Y} \cdot \mathbf{u}_2)^2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_d)^2 \\
&= Y_1^2 + Y_2^2 + \dots + Y_d^2,
\end{aligned}$$

$$\begin{aligned}
E(\|P_W \mathbf{Y}\|^2) &= E(Y_1^2 + Y_2^2 + \dots + Y_d^2) \\
&= E(Y_1^2) + E(Y_2^2) + \dots + E(Y_d^2) \\
&= \sum_{i=1}^d E(Y_i^2).
\end{aligned}$$

Como $\text{var}(Y_i) = E(Y_i^2) - (E(Y_i))^2$, temos que:

$$\begin{aligned}
E(\|P_W \mathbf{Y}\|^2) &= \sum_{i=1}^d (\text{var}(Y_i) + (E(Y_i))^2) \\
&= \sum_{i=1}^d (\sigma^2 + \tau_i^2) \\
&= d\sigma^2 + \sum_{i=1}^d \tau_i^2 \\
&= d\sigma^2 + \|P_W \boldsymbol{\tau}\|^2.
\end{aligned}$$

Uma vez que

$$\begin{aligned}
\|P_W \boldsymbol{\tau}\|^2 &= (\boldsymbol{\tau} \cdot \mathbf{u}_1)^2 + (\boldsymbol{\tau} \cdot \mathbf{u}_2)^2 + \dots + (\boldsymbol{\tau} \cdot \mathbf{u}_d)^2 \\
&= \tau_1^2 + \tau_2^2 + \dots + \tau_d^2 \\
&= \sum_{i=1}^d \tau_i^2.
\end{aligned}$$

2.4 Análise de variância

A análise de variância é uma técnica utilizada para analisar variabilidade nos dados. Caso o pesquisador deseje montar um experimento e analisar os dados, por meio dessa técnica se faz necessário garantir alguns princípios básicos, os quais são: repetição, aleatorização e controle local.

A repetição dos tratamentos no experimento é necessária, porque sem ela não existe estimador da variância. A aleatorização garante que nenhum dos tratamentos está sendo beneficiado no planejamento do experimento. O controle local é necessário apenas quando a área experimental não é totalmente homogênea, de modo que este consiste no arranjo de unidades experimentais em blocos que são similares entre si. O uso do controle local caracteriza diferentes tipos de delineamentos experimentais.

Cada delineamento experimental resulta em decomposições distintas de V em soma direta ⁸ de subespaços. A partir da decomposição de V é possível obter o esquema da análise de variância onde cada subespaço constitui uma fonte de variação. As fontes de variação são essenciais para testar hipóteses de interesse no experimento.

A decomposição de V leva à decomposição do vetor de respostas \mathbf{Y} , por meio dos vetores de projeção de \mathbf{Y} nos subespaços de V . Como os vetores de projeção são vetores aleatórios, é possível realizar inferência. Por meio dos vetores de projeção, é possível definir alguns conceitos fundamentais na análise de variância.

Definição 2.4.1. *Seja W um subespaço de V . A soma de quadrados de \mathbf{Y} para W é dada por $\|P_W \mathbf{Y}\|^2$. O número de graus de liberdade para W é dado por $\dim(W)$. O quadrado médio para W é dado por:*

$$QM(W) = \frac{\|P_W \mathbf{Y}\|^2}{\dim(W)}.$$

Definição 2.4.2. *A esperança do quadrado médio para W é dada por:*

$$EQM(W) = \frac{E(\|P_W \mathbf{Y}\|^2)}{\dim(W)}.$$

Definidos esses termos, é possível se obter a análise de variância para os diferentes tipos de delineamentos, que são apresentados, a seguir, no Capítulo 3.

⁸ Veja Definição .0.7 do Apêndice deste trabalho.

3 O MÉTODO GEOMÉTRICO PARA ANÁLISE DE VARIÂNCIA EM DELINEAMENTOS EXPERIMENTAIS

Neste capítulo, são apresentados os tipos de delineamentos experimentais, que resultam em diferentes tabelas de análise de variância. Neste capítulo, a análise de variância é apresentada com enfoque em conceitos geométricos.

3.1 Delineamento inteiramente casualizado

No delineamento inteiramente casualizado, as parcelas do experimento são totalmente homogêneas, ou seja, possuem condições iguais. Nesse caso, os tratamentos podem ser alocados aleatoriamente às parcelas, por sorteio. O que significa que não existe estrutura de blocos nas parcelas.

Para esse caso particular de delineamento, o espaço vetorial V pode ser decomposto em soma de três subespaços. A análise de variância pode ser obtida, por meio das projeções ortogonais de \mathbf{Y} nesses subespaços de V . Para tanto, é necessário definir esses subespaços.

O subespaço V_T

O subespaço V_T , também chamado de subespaço de tratamento, consiste nos vetores em V que são constantes em cada tratamento. Para se obter a projeção ortogonal de \mathbf{Y} em V_T é necessário obter uma base que gere esse subespaço. Tal base é obtida da seguinte maneira:

Para cada tratamento i , \mathbf{u}_i é o vetor com coordenadas iguais a:

$$\begin{cases} 1, & \text{se } T(\omega) = i, \\ 0, & \text{caso contrário.} \end{cases}$$

Exemplo

Suponha um experimento fictício em que existem três tratamentos A, B, e C com $r_A = 5$, $r_B = r_C = 4$ e $N = 13$. Na Tabela 3.1 mostra-se o conjunto Ω de todas as parcelas, o conjunto \mathcal{T} de tratamentos, o vetor de respostas \mathbf{Y} , os vetores \mathbf{u}_A , \mathbf{u}_B e \mathbf{u}_C da base ortogonal para V_T , o vetor de médias $\boldsymbol{\tau}$ e o vetor de médias ajustadas $\hat{\boldsymbol{\tau}}$. Note que, nesse exemplo, $\boldsymbol{\tau}$ e $\hat{\boldsymbol{\tau}}$ são combinações lineares dos vetores da base.

Teorema 3.1.1. Para cada i , $\mathbf{u}_i \cdot \mathbf{u}_i = r_i$.

Tabela 3.1 – Vetores no experimento inteiramente casualizado.

Ω	\mathcal{T}	\mathbf{Y}	\mathbf{u}_A	\mathbf{u}_B	\mathbf{u}_C	$\boldsymbol{\tau}$	$\hat{\boldsymbol{\tau}}$
1	B	Y_1	0	1	0	τ_B	$\hat{\tau}_B$
2	A	Y_2	1	0	0	τ_A	$\hat{\tau}_A$
3	C	Y_3	0	0	1	τ_C	$\hat{\tau}_C$
4	C	Y_4	0	0	1	τ_C	$\hat{\tau}_C$
5	B	Y_5	0	1	0	τ_B	$\hat{\tau}_B$
6	A	Y_6	1	0	0	τ_A	$\hat{\tau}_A$
7	B	Y_7	0	1	0	τ_B	$\hat{\tau}_B$
8	A	Y_8	1	0	0	τ_A	$\hat{\tau}_A$
9	C	Y_9	0	0	1	τ_C	$\hat{\tau}_C$
10	C	Y_{10}	0	0	1	τ_C	$\hat{\tau}_C$
11	A	Y_{11}	1	0	0	τ_A	$\hat{\tau}_A$
12	B	Y_{12}	0	1	0	τ_B	$\hat{\tau}_B$
13	A	Y_{13}	1	0	0	τ_A	$\hat{\tau}_A$

Fonte: Adaptado Bailey (2008).

Demonstração. O produto interno $\mathbf{u}_i \cdot \mathbf{u}_i$ é definido por:

$$\begin{aligned}
 \mathbf{u}_i \cdot \mathbf{u}_i &= \sum_{\omega \in \Omega} (u_i)_\omega (u_i)_\omega \\
 &= \sum_{\substack{\omega \in \Omega, \\ T(\omega)=i}} 1 \times 1 \\
 &= r_i.
 \end{aligned}$$

Como na parcela ω onde se encontra o tratamento i atribuímos o valor 1, e nas demais 0, essa soma de quadrados resulta no número de repetições do tratamento i , ou seja, r_i . \square

Como já sabemos, a dimensão de um espaço vetorial é definida pelo número de vetores da base. Note que o conjunto de vetores $\{\mathbf{u}_1, \dots, \mathbf{u}_t\}$ forma uma base V_T , pois qualquer vetor \mathbf{v} em V_T é uma combinação linear única desses vetores. Logo, se existem t tratamentos tem-se consequentemente, t vetores na base, e portanto $\dim(V_T) = t$.

Utilizando-se o Teorema 2.2.3, é possível obter-se a projeção de \mathbf{Y} em V_T . Para tanto, seja $\{\mathbf{u}_1, \dots, \mathbf{u}_t\}$ uma base para V_T , então a projeção ortogonal é dada por:

$$\begin{aligned}
 P_{V_T} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_t}{\mathbf{u}_t \cdot \mathbf{u}_t} \right) \mathbf{u}_t \\
 &= \sum_{i=1}^t \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i.
 \end{aligned}$$

Note que $\mathbf{Y} \cdot \mathbf{u}_i$ é a soma dos valores das coordenadas Y_i de \mathbf{Y} para o tratamento i , em que i significa a posição relativa a uma parcela ω onde se alocou o tratamento i e r_i o número de vezes que esse tratamento foi replicado. Será utilizada a notação:

$$\mathbf{Y} \cdot \mathbf{u}_i = \sum_{i=1}^{r_i} Y_i = \text{SOMA}_{T=i}.$$

Dessa forma, a projeção de \mathbf{Y} em V_T é:

$$\begin{aligned} P_{V_T} \mathbf{Y} &= \sum_{i=1}^t \left(\frac{\text{SOMA}_{T=i}}{r_i} \right) \mathbf{u}_i \\ &= \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i \\ &= \text{vetor de médias dos tratamentos.} \end{aligned}$$

O subespaço V_0

Seja \mathbf{u}_0 o vetor definido por $\mathbf{u}_0 = (1, 1, \dots, 1)$. O subespaço V_0 , também chamado de subespaço de média geral, consiste nos múltiplos escalares de \mathbf{u}_0 . A base de V_0 é $\{\mathbf{u}_0\}$ e, portanto, $\dim(V_0) = 1$.

A projeção de \mathbf{Y} em V_0 é obtida pelo Teorema 2.2.3 como segue:

$$P_{V_0} \mathbf{Y} = \left(\frac{\mathbf{Y} \cdot \mathbf{u}_0}{\mathbf{u}_0 \cdot \mathbf{u}_0} \right) \mathbf{u}_0.$$

Note que $\mathbf{Y} \cdot \mathbf{u}_0$ é a soma de todos os valores Y_i de \mathbf{Y} ; será utilizada a notação:

$$\mathbf{Y} \cdot \mathbf{u}_0 = \sum_{i=1}^N Y_i = \text{SOMA}_Y.$$

Dessa forma, a projeção de \mathbf{Y} em V_0 é:

$$\begin{aligned} P_{V_0} \mathbf{Y} &= \left(\frac{\text{SOMA}_Y}{N} \right) \mathbf{u}_0 \\ &= \bar{Y} \mathbf{u}_0 \\ &= \text{vetor de média geral.} \end{aligned}$$

O subespaço W_T

O subespaço W_T , também chamado de subespaço de efeito de tratatamento, é dado pelo conjunto de vetores em V_T que são perpendiculares a V_0 , e pode ser definido como segue:

$$W_T = V_T \cap V_0^\perp = \{\mathbf{v} \in V_T : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

Perceba que após definir V_0 e W_T , o subespaço V_T foi decomposto em soma direta desses subespaços, ou seja:

$$V_T = V_0 \oplus W_T.$$

Desse modo, $\dim(V_T) = \dim(V_0) + \dim(W_T)$, de onde vem que:

$$\begin{aligned} \dim(W_T) &= \dim(V_T) - \dim(V_0) \\ &= t - 1. \end{aligned}$$

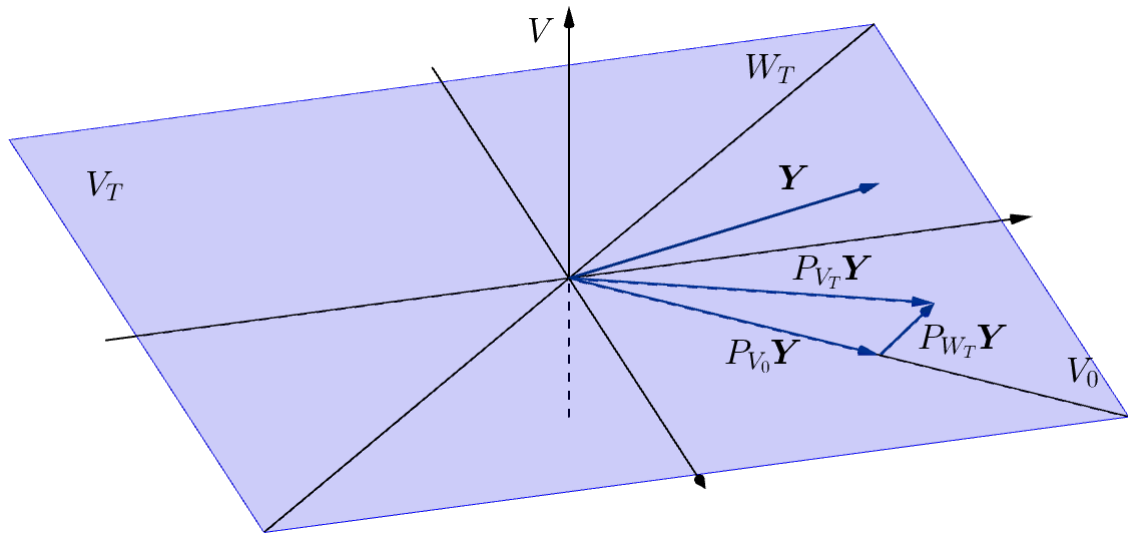
A decomposição de V_T leva também à decomposição de vetores, de modo que, qualquer vetor em V_T pode ser decomposto como soma de vetores em V_0 e W_T . Sendo assim, $P_{V_T}\mathbf{Y} = P_{V_0}\mathbf{Y} + P_{W_T}\mathbf{Y}$. Logo, a projeção de \mathbf{Y} em W_T pode ser obtida por diferença:

$$\begin{aligned} P_{W_T}\mathbf{Y} &= P_{V_T}\mathbf{Y} - P_{V_0}\mathbf{Y} \\ &= \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \bar{Y} \mathbf{u}_0 \\ &= \text{vetor de médias dos tratamentos} - \text{vetor de média geral} \\ &= \text{vetor de efeitos dos tratamentos} . \end{aligned}$$

O espaço vetorial V , os subespaços V_T , V_0 e W_T e as projeções de \mathbf{Y} nesses subespaços estão representados na Figura 3.1.

O subespaço V_T^\perp

O subespaço V_T^\perp , também chamado de subespaço de resíduo, consiste no complemento ortogonal de V_T . Esse conjunto é constituído por todos os vetores em V que são perpendiculares

Figura 3.1 – Projeções de \mathbf{Y} em V_0 e W_T .

Fonte: Da autora (2019).

a V_T , e é definido como:

$$V_T^\perp = \{\mathbf{v} \in V : \mathbf{v} \text{ é ortogonal a } \mathbf{w} \text{ para todo } \mathbf{w} \in V_T\}.$$

Agora o espaço vetorial V foi decomposto como soma direta dos subespaços:

$$V = V_T \oplus V_T^\perp.$$

Desse modo, a dimensão de V corresponde à soma das dimensões de:

$$\dim(V) = \dim(V_T) + \dim(V_T^\perp).$$

Assim, a dimensão de V_T^\perp pode ser obtida por diferença:

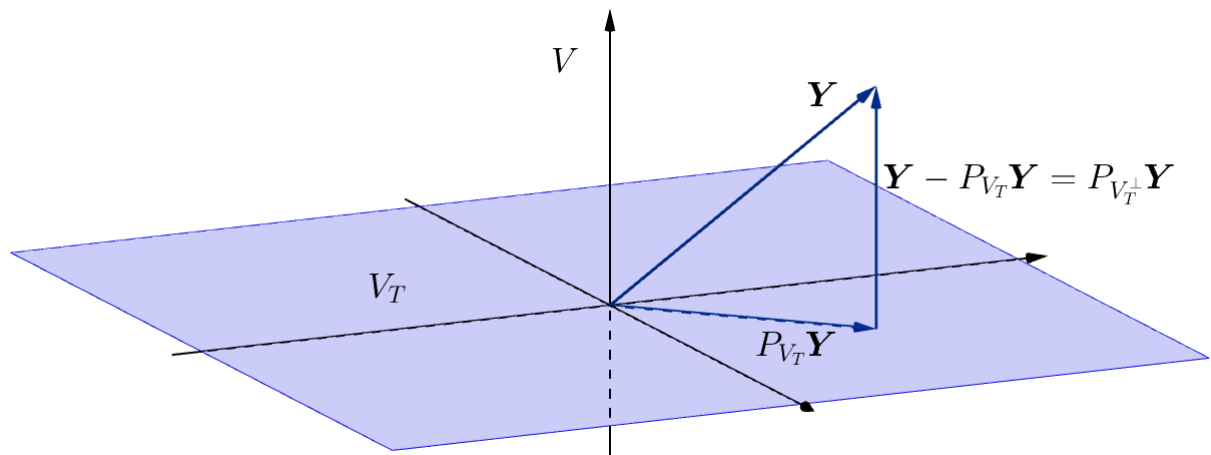
$$\begin{aligned} \dim(V_T^\perp) &= \dim(V) - \dim(V_T) \\ &= N - t. \end{aligned}$$

A projeção ortogonal de \mathbf{Y} em V_T^\perp também pode ser obtida pela diferença de \mathbf{Y} e $P_{V_T}\mathbf{Y}$ como segue:

$$\begin{aligned} P_{V_T^\perp}\mathbf{Y} &= \mathbf{Y} - P_{V_T}\mathbf{Y} \\ &= \mathbf{Y} - \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i \\ &= \text{vetor de dados} - \text{vetor de médias dos tratamentos} \\ &= \text{vetor de resíduos.} \end{aligned}$$

O espaço vetorial V e a projeção de \mathbf{Y} em V_T^\perp estão representadas na Figura 3.2.

Figura 3.2 – Projeção de \mathbf{Y} em V_T^\perp .



Fonte: Da autora (2019).

Como $V_T = V_0 \oplus W_T$ temos que:

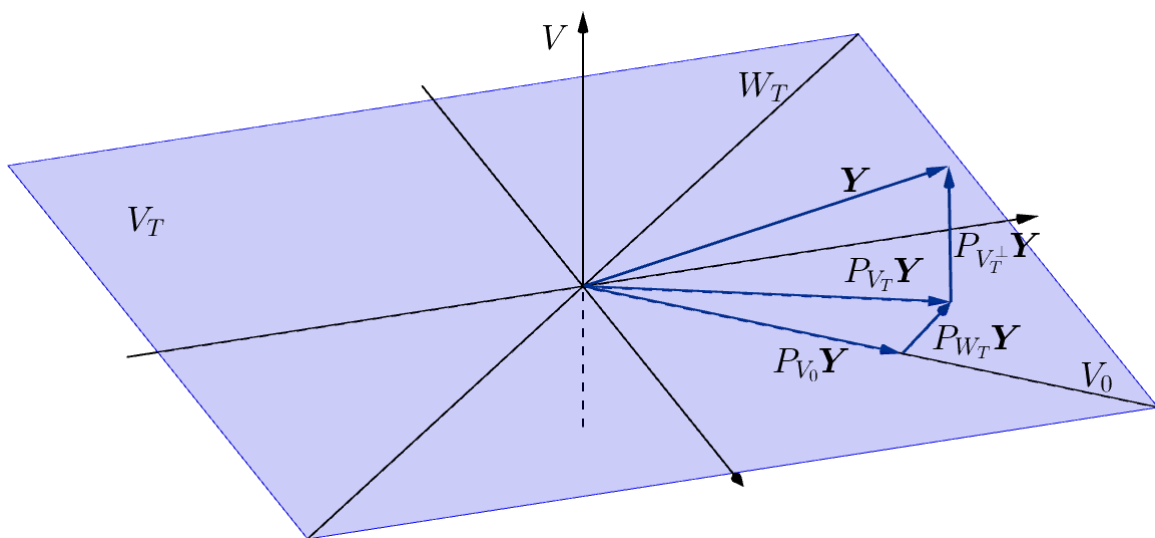
$$V = V_0 \oplus W_T \oplus V_T^\perp.$$

Essa decomposição de V é única, o que significa que todo vetor em V se escreve de modo único como soma de vetores nos subespaços V_0 , W_T e V_T^\perp . Desse modo, o vetor \mathbf{Y} pode ser decomposto em soma de três vetores (Figura 3.3). Note que $\dim(V) = N$ corresponde à

soma das dimensões desses subespaços.

$$\begin{aligned}
 \dim(V) &= \dim(V_0) + \dim(W_T) + \dim(V_T^\perp) \\
 &= 1 + (\dim(V_T) - \dim(V_0)) + (\dim(V) - \dim(V_T)) \\
 &= 1 + (t - 1) + (N - t) \\
 &= N.
 \end{aligned}$$

Figura 3.3 – Decomposição de \mathbf{Y} .



Fonte: Da autora (2019).

Agora é possível usar a Definição 2.4.1 para obter-se as somas de quadrados, os graus de liberdade e os quadrados médios para os subespaços definidos anteriormente. Para isso, basta obter a norma ao quadrado dos vetores de projeção e dividir pela dimensão dos respectivos subespaços.

A soma de quadrados para V_0 , também chamada de soma de quadrados para média geral é dada por:

$$\begin{aligned}
 \|P_{V_0}\mathbf{Y}\|^2 &= P_{V_0}\mathbf{Y} \cdot P_{V_0}\mathbf{Y} \\
 &= \left(\frac{\text{SOMA}_Y}{N}\right) \mathbf{u}_0 \cdot \left(\frac{\text{SOMA}_Y}{N}\right) \mathbf{u}_0 \\
 &= \left(\frac{\text{SOMA}_Y}{N}\right)^2 \mathbf{u}_0 \cdot \mathbf{u}_0 \\
 &= \frac{(\text{SOMA}_Y)^2}{N^2} N \\
 &= \frac{(\text{SOMA}_Y)^2}{N}.
 \end{aligned}$$

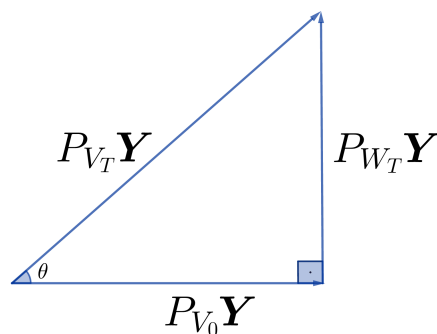
O quadrado médio para V_0 é dado por:

$$QM(V_0) = \frac{\|P_{V_0}\mathbf{Y}\|^2}{\dim(V_0)} = \frac{\|P_{V_0}\mathbf{Y}\|^2}{1}.$$

A soma de quadrados para W_T , também chamada de soma de quadrados dos efeitos dos tratamentos, pode ser obtida por diferença, veja Figura 3.4. Usando o teorema de Pitágoras temos que $P_{V_T}\mathbf{Y}$ é soma de vetores ortogonais $P_{V_0}\mathbf{Y}$ e $P_{W_T}\mathbf{Y}$, sendo assim:

$$\begin{aligned}
 \|P_{V_T}\mathbf{Y}\|^2 &= \|P_{V_0}\mathbf{Y}\|^2 + \|P_{W_T}\mathbf{Y}\|^2 \\
 \|P_{W_T}\mathbf{Y}\|^2 &= \|P_{V_T}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2.
 \end{aligned}$$

Figura 3.4 – Pitágoras.



Fonte: Da autora (2019).

Em que $\|P_{V_T}\mathbf{Y}\|^2$ é a soma de quadrados para V_T , também chamada de soma de quadrados das médias dos tratamentos, dada por:

$$\begin{aligned}\|P_{V_T}\mathbf{Y}\|^2 &= P_{V_T}\mathbf{Y} \cdot P_{V_T}\mathbf{Y} \\ &= \sum_{i=1}^t \left(\frac{\text{SOMA}_{T=i}}{r_i} \right) \mathbf{u}_i \cdot \sum_{i=1}^t \left(\frac{\text{SOMA}_{T=i}}{r_i} \right) \mathbf{u}_i \\ &= \sum_{i=1}^t \left(\frac{\text{SOMA}_{T=i}}{r_i} \right)^2 \mathbf{u}_i \cdot \mathbf{u}_i \\ &= \sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i^2} r_i \\ &= \sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i}.\end{aligned}$$

O quadrado médio para W_T é dado por:

$$QM(W_T) = \frac{\|P_{W_T}\mathbf{Y}\|^2}{\dim(W_T)} = \frac{\|P_{W_T}\mathbf{Y}\|^2}{t-1}.$$

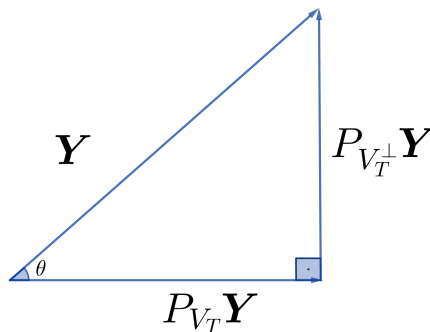
A soma de quadrados para V_T^\perp , também chamada de soma de quadrados dos resíduos, também pode ser obtida por diferença, veja Figura 3.5. Novamente por Pitágoras \mathbf{Y} é soma dos vetores ortogonais $P_{V_T}\mathbf{Y}$ e $P_{V_T^\perp}\mathbf{Y}$, de onde vem que:

$$\|\mathbf{Y}\|^2 = \|P_{V_T}\mathbf{Y}\|^2 + \|P_{V_T^\perp}\mathbf{Y}\|^2$$

$$\|P_{V_T^\perp}\mathbf{Y}\|^2 = \|\mathbf{Y}\|^2 - \|P_{V_T}\mathbf{Y}\|^2,$$

em que $\|\mathbf{Y}\|^2 = \sum_{\omega \in \Omega} Y_\omega^2$ é a soma total de quadrados.

Figura 3.5 – Pitágoras.



Fonte: Da autora (2019).

O quadrado médio para V_T^\perp é dado por:

$$QM(V_T^\perp) = \frac{\|P_{V_T^\perp} \mathbf{Y}\|^2}{\dim(V_T^\perp)} = \frac{\|P_{V_T^\perp} \mathbf{Y}\|^2}{N-t}.$$

Agora, é possível, obter-se a análise de variância para o delineamento inteiramente casualizado, como mostra a Tabela 3.2. Note que as fontes de variação correspondem aos subespaços da decomposição de V como soma direta.

Tabela 3.2 – Tabela de análise de variância para delineamento inteiramente casualizado.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0} \mathbf{Y}\ ^2$	$QM(V_0)$	$QM(V_0)/QM(V_T^\perp)$
Tratamentos W_T	$t-1$	$\ P_{W_T} \mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/QM(V_T^\perp)$
Resíduo V_T^\perp	$N-t$	$\ P_{V_T^\perp} \mathbf{Y}\ ^2$	$QM(V_T^\perp)$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

Geralmente, os livros didáticos não trazem o ajuste V_0 na análise de variância. Porém, de acordo com Bailey (2008), os cálculos são mais claros, quando as somas de quadrados e graus de liberdade somam os totais gerais de \mathbf{Y} e não os totais sem o ajuste para V_0 . Assim, na análise de variância sem a linha para a média, o total geral é corrigido para a média, ou seja, Soma de Quadrados Totais = $\|\mathbf{Y}\|^2 - \|P_{V_0} \mathbf{Y}\|^2$.

Como aponta Adão (2011), os delineamentos são utilizados para validar estatisticamente hipóteses sobre efeitos de tratamentos. As hipóteses podem ser testadas pela razão de variâncias que é a razão entre dois quadrados médios.

Em análise de variância, para realizar um teste é necessário assumir alguns pressupostos, os quais são:

- Aditividade (o vetor de respostas é decomposto em soma de vetores relativos à decomposição do espaço de dados em soma direta de subespaços);
- Normalidade (o vetor de dados segue distribuição normal de probabilidade);
- Independência (o vetor de resíduos é ortogonal aos demais vetores, ou seja, não existe correlação entre esses vetores);
- Homogeneidade de variâncias (somadas de quadrados possuem variâncias comuns).

Sob as suposições acima, como veremos a seguir, é possível realizar o teste F na análise de variância. Desse modo, assumamos que $\mathbf{Y} \sim N(\boldsymbol{\tau}, \sigma^2 \mathbf{I})$ onde $\boldsymbol{\tau} \in V_T$, de modo que o modelo aditivo para soma de vetores é correto e que qualquer variação no experimento é decorrente do erro. Em termos vetoriais temos:

$$\mathbf{Y} = \boldsymbol{\tau} + \boldsymbol{\varepsilon},$$

em que $E(\mathbf{Y}) = \boldsymbol{\tau}$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ e $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$.

Note que, $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_T$ com $\boldsymbol{\tau}_0 \in V_0$ e $\boldsymbol{\tau}_T \in W_T$. Dessa forma, é possível estimar combinações lineares dos efeitos dos tratamentos, isso porque, $\boldsymbol{\tau}_T = \boldsymbol{\tau} - \boldsymbol{\tau}_0$.

O teste de hipótese

Na Tabela 3.2, tem-se a coluna referente à estatística F , que é dada pela razão de variâncias. Como já mencionado, em análise de variância, o objetivo é realizar um teste sobre os efeitos dos tratamentos utilizados no experimento. Para realizar esse teste, é necessário conhecer a distribuição da razão de variâncias sob a hipótese nula. O teorema, a seguir, mostra essa razão possui distribuição F .

Teorema 3.1.2. *Suponha que \mathbf{Y} tem distribuição normal multivariada, $E(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$ e $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$. Sejam W_1 e W_2 subespaços de V com dimensões d_1 e d_2 respectivamente. Então, temos o seguinte:*

- i) *Se $P_{W_1} \boldsymbol{\tau} = \mathbf{0}$ então $\|P_{W_1} \mathbf{Y}\|^2 / \sigma^2$ tem distribuição χ^2 com d_1 graus de liberdade.*
- ii) *Se W_1 é ortogonal a W_2 e $P_{W_1} \boldsymbol{\tau} = P_{W_2} \boldsymbol{\tau} = \mathbf{0}$ então a relação $QM(W_1) / QM(W_2)$ tem distribuição F com d_1 e d_2 graus de liberdade.*

Demonstração.

- i) Seja $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{d_1}\}$ uma base ortonormal para W_1 . Expressando $P_{W_1} \mathbf{Y}$ com relação a essa base temos:

$$\begin{aligned} P_{W_1} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_{d_1}}{\mathbf{u}_1 \cdot \mathbf{u}_{d_1}} \right) \mathbf{u}_{d_1} \\ &= (\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1}) \mathbf{u}_{d_1}. \end{aligned}$$

Calculando a esperança de $\mathbf{Y} \cdot \mathbf{u}_i$ e usando o fato de que $P_{W_1} \boldsymbol{\tau} = \mathbf{0}$ temos:

$$E(\mathbf{Y} \cdot \mathbf{u}_i) = E(\mathbf{Y}) \cdot \mathbf{u}_i = \boldsymbol{\tau} \cdot \mathbf{u}_i = 0.$$

Calculando a variância de $\mathbf{Y} \cdot \mathbf{u}_i$ e usando o Teorema 3.1.3 que será visto mais adiante, temos que:

$$\text{var}(\mathbf{Y} \cdot \mathbf{u}_i) = \text{var}(\mathbf{u}_i \cdot \mathbf{Y}) = \|\mathbf{u}_i\|^2 \sigma^2 = \sigma^2.$$

Logo, $\mathbf{Y} \cdot \mathbf{u}_i$ são variáveis normais independentes com média 0 e variância σ^2 . Portanto segue que:

$$\begin{aligned} \frac{\|P_{W_1} \mathbf{Y}\|^2}{\sigma^2} &= \frac{P_{W_1} \mathbf{Y} \cdot P_{W_1} \mathbf{Y}}{\sigma^2} \\ &= \frac{1}{\sigma^2} ((\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1}) \mathbf{u}_{d_1}) \cdot \\ &\quad ((\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1}) \mathbf{u}_{d_1}) \\ &= \frac{1}{\sigma^2} ((\mathbf{Y} \cdot \mathbf{u}_1)^2 + (\mathbf{Y} \cdot \mathbf{u}_2)^2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1})^2) \\ &= \frac{1}{\sigma^2} \left(\sum_{i=1}^{d_1} (\mathbf{Y} \cdot \mathbf{u}_i)^2 \right) \\ &= \sum_{i=1}^{d_1} \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sigma} \right)^2. \end{aligned}$$

Como $\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sigma}$ tem distribuição normal padrão, segue que $\sum_{i=1}^{d_1} \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sigma} \right)^2$ tem distribuição χ^2 com d_1 graus de liberdade.

ii) Como W_1 e W_2 são subespaços ortogonais, $P_{W_1} \mathbf{Y}$ e $P_{W_2} \mathbf{Y}$ são vetores aleatórios ortogonais independentes¹, de onde vem que:

$$\frac{\|P_{W_1} \mathbf{Y}\|^2}{\sigma^2} \text{ e } \frac{\|P_{W_2} \mathbf{Y}\|^2}{\sigma^2}$$

tem distribuição χ^2 com d_1 e d_2 graus de liberdade, respectivamente. Logo, segue que:

$$\frac{\|P_{W_1} \mathbf{Y}\|^2 / \sigma^2}{\|P_{W_2} \mathbf{Y}\|^2 / \sigma^2} = \frac{\|P_{W_1} \mathbf{Y}\|^2}{\|P_{W_2} \mathbf{Y}\|^2},$$

¹ A independência entre esses vetores será vista no Capítulo 5, Teorema 5.1.1, item iv), item b).

tem distribuição F com d_1 e d_2 graus de liberdade. Portanto,

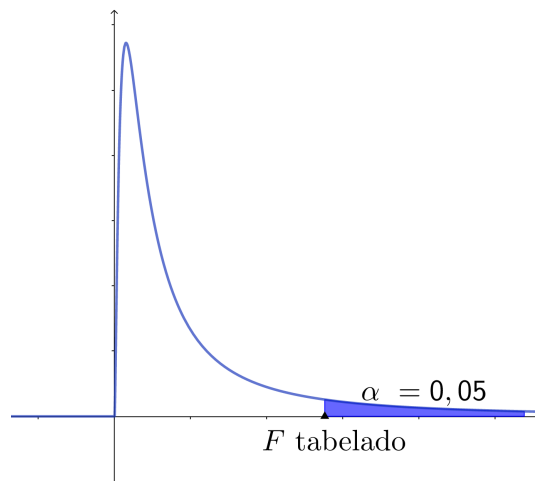
$$\frac{\|P_{W_1}\mathbf{Y}\|^2/d_1}{\|P_{W_2}\mathbf{Y}\|^2/d_2} = \frac{QM(W_1)}{QM(W_2)}$$

tem distribuição F com d_1 e d_2 graus de liberdade.

□

No Teorema 3.1.2 mostra-se que a razão de variâncias $QM(W_1)/QM(W_2)$ tem distribuição F com d_1 e d_2 graus de liberdade. De acordo com Casella e Berger (2010), a distribuição F surge naturalmente como distribuição da razão de variâncias. Essa razão é utilizada para realizar o teste de hipótese na análise de variância. Note que, esse teste é unilateral à direita, veja Figura 3.6.

Figura 3.6 – Distribuição F .



Fonte: Da autora (2019).

Para inferir conclusões na análise de variância é preciso seguir os seguintes passos:

- 1) Estabeleça a hipótese nula (H_0) e alternativa (H_a) a serem testadas;
- 2) Fixe o nível de significância α (geralmente 0,01 ou 0,05);
- 3) Determine a região crítica de rejeição de H_0 , para tanto, obtenha o valor F tabelado (F_t) para d_1 e d_2 graus de liberdade;
- 4) Obtenha o valor F calculado (F_c) dado pela razão de variação;
- 5) Compare F_c com F_t , se $F_c > F_t$ rejeita-se H_0 .

Agora é possível utilizar o Teorema 3.1.2, para realizar o teste F para razão de variâncias, para o delineamento inteiramente casualizado.

Para testar se a média geral do experimento é zero, pode-se testar a hipótese nula H_0 , contra a hipótese alternativa H_a , as quais são:

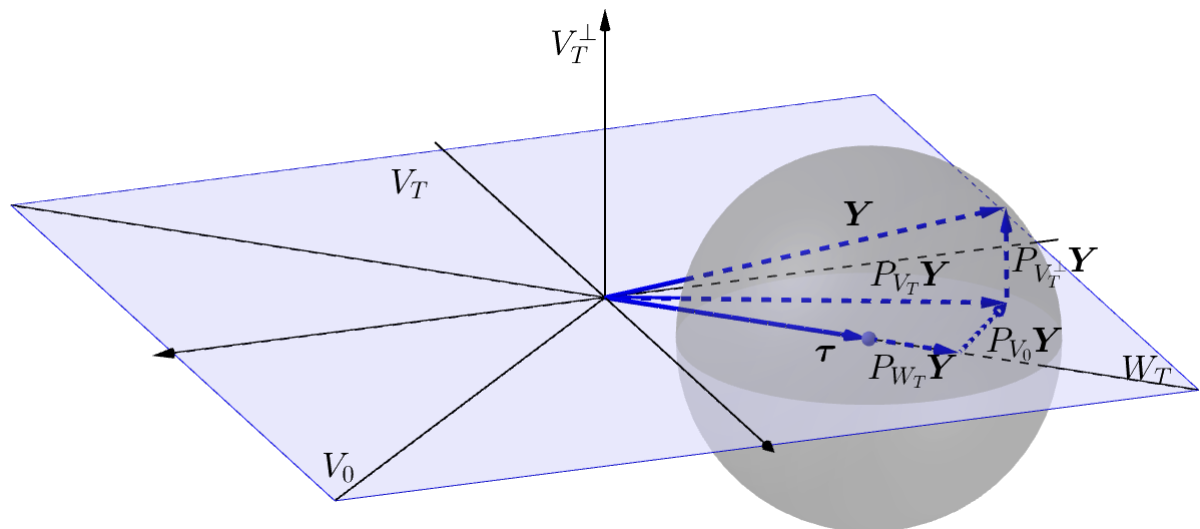
$$H_0 : \bar{\tau} = 0$$

$$H_a : \bar{\tau} \neq 0.$$

Se a hipótese nula é verdadeira, então $\sum_i \tau_i = 0$, logo, $\boldsymbol{\tau} \in W_T$. Se a hipótese nula é falsa, então $\sum_i \tau_i \neq 0$, logo, $\boldsymbol{\tau} \notin W_T$. Sob H_0 verdadeira, a nuvem de pontos é centrada em $\boldsymbol{\tau}$ que pertence ao subespaço W_T , como mostra a Figura 3.7.

Para realizar o teste, a ideia é comparar o vetor $P_{V_0}\mathbf{Y}$ com o vetor $P_{V_T^\perp}\mathbf{Y}$ (Figura 3.7). Isso porque o vetor de resíduos detecta possíveis variações no experimento. Para comparar esses vetores, use a razão entre $QM(V_0)$ e $QM(V_T^\perp)$. Se essa razão for menor que um, não se rejeita H_0 e pode-se concluir que a média geral do experimento pode ser zero. Caso contrário, se essa razão for maior que um, rejeita-se H_0 e pode-se concluir que a média geral do experimento é diferente de zero.

Figura 3.7 – Teste de hipótese para média geral.



Fonte: Da autora (2019).

Entretanto, essa hipótese não é interessante, pois ela não indica se existe diferença entre as médias dos tratamentos.

Para testar se existe diferença entre as médias dos tratamentos, pode-se testar a hipótese nula H_0 contra a hipótese alternativa H_a , as quais são:

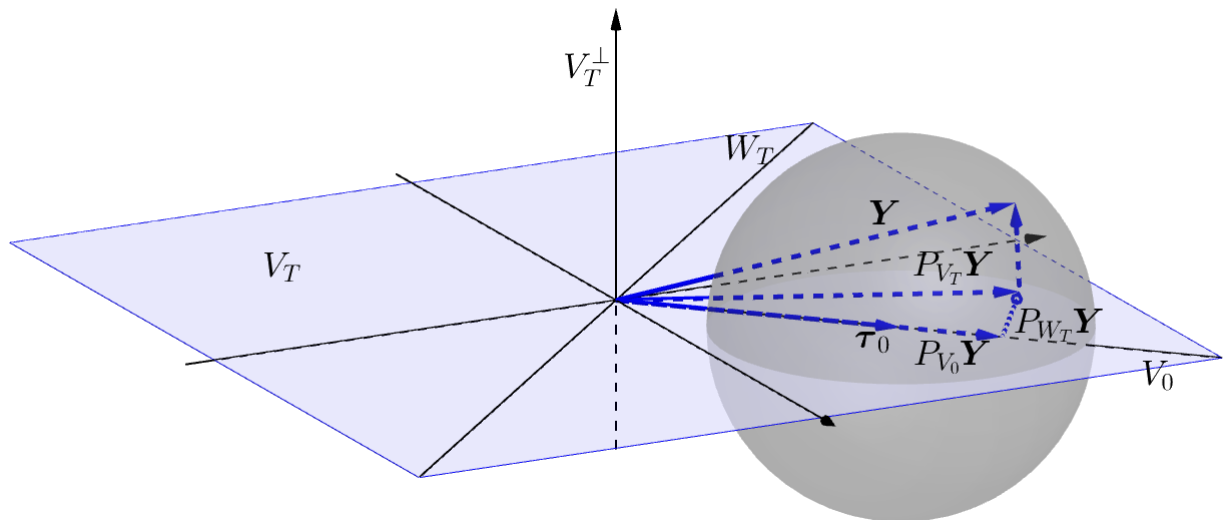
$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t$$

$H_a : \tau$ não é um vetor constante .

Se a hipótese nula é verdadeira τ é um vetor constante, logo, $\tau \in V_0$, e portanto $\tau = \tau_0$. Se a hipótese nula é falsa τ não é um vetor constante, logo, $\tau = \tau_0 + \tau_T$ e portanto existe efeito de tratamento. Sob H_0 verdadeira, a nuvem de pontos é centrada em τ_0 que pertence ao subespaço V_0 , como mostra a Figura 3.8.

Para realizar o teste, a idéia é comparar o vetor $P_{W_T}Y$ com o vetor de resíduos $P_{V_T^\perp}Y$ (Figura 3.8). Isso porque, o vetor de resíduos detecta possíveis variações do experimento. Para comparar esses vetores, use a razão entre $QM(W_T)$ e $QM(V_T^\perp)$. Se essa razão for menor que um, não se rejeita H_0 e pode-se concluir que não existe diferença entre as médias dos tratamentos. Caso contrário, se essa razão for maior que um, rejeita-se H_0 e pode-se concluir que existe diferença entre as médias dos tratamentos.

Figura 3.8 – Teste de hipótese para tratamentos.



Fonte: Da autora (2019).

Vale destacar que, o teste descrito indica se existe ou não diferença entre as médias dos tratamentos. Entretanto, não indica qual efeito de tratamento é estatisticamente melhor. Para tanto, é necessário fazer um teste de comparação múltipla.

Estimadores

Para variáveis aleatórias independentes, a correlação entre pares diferentes de variáveis aleatórias é nula, então a matriz de variâncias e covariâncias é $\sigma^2 I$. Nesse caso, temos o resultado que segue.

Teorema 3.1.3. Assuma que $E(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$ e $\text{cov}(\mathbf{Y}) = \sigma^2 \mathbf{I}$. Sejam \mathbf{x} e \mathbf{z} vetores em V_T . Então

i) A variância do estimador $\mathbf{x} \cdot \mathbf{Y}$ é $\|\mathbf{x}\|^2 \sigma^2$;

ii) A covariância de $\mathbf{x} \cdot \mathbf{Y}$ e $\mathbf{z} \cdot \mathbf{Y}$ é $(\mathbf{x} \cdot \mathbf{z}) \sigma^2$.

iii) O melhor (isto é, de variância mínima) estimador linear não viesado do escalar $\mathbf{x} \cdot \boldsymbol{\tau}$ é $\mathbf{x} \cdot \mathbf{Y}$.

Demonstração.

i) Primeiramente vamos reescrever $(\mathbf{x} \cdot \mathbf{Y})^2$ como forma quadrática para o caso $N = 2$.

$$\begin{aligned} (\mathbf{x} \cdot \mathbf{Y})^2 &= (\mathbf{x} \cdot \mathbf{Y})(\mathbf{x} \cdot \mathbf{Y}) \\ &= (x_1 Y_1 + x_2 Y_2)(x_1 Y_1 + x_2 Y_2) \\ &= x_1^2 Y_1^2 + 2x_1 Y_1 x_2 Y_2 + x_2^2 Y_2^2 \\ &= \begin{pmatrix} Y_1 & Y_2 \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 x_2 \\ x_2 x_1 & x_2^2 \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \end{pmatrix} \\ &= \mathbf{Y}^\top \mathbf{A} \mathbf{Y}. \end{aligned}$$

Para o caso, geral temos que:

$$(\mathbf{x} \cdot \mathbf{Y})^2 = \begin{pmatrix} Y_1 & \dots & Y_N \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 x_2 & \dots & x_1 x_N \\ x_2 x_1 & x_2^2 & \dots & x_2 x_N \\ \dots & \dots & \dots & \dots \\ x_N x_1 & x_N x_2 & \dots & x_N^2 \end{pmatrix} \begin{pmatrix} Y_1 \\ \dots \\ Y_N \end{pmatrix} = \mathbf{Y}^\top \mathbf{A} \mathbf{Y}.$$

Reescrevendo $(\mathbf{x} \cdot \mathbf{Y})^2$ como forma quadrática, é possível utilizar o Teorema .0.1 que encontra-se no Apêndice deste trabalho para demonstrar o resultado desejado.

Calculando agora a variância do estimador $\mathbf{x} \cdot \mathbf{Y}$ obtemos:

$$\begin{aligned}
 \text{var}(\mathbf{x} \cdot \mathbf{Y}) &= E(\mathbf{x} \cdot \mathbf{Y})^2 - (E(\mathbf{x} \cdot \mathbf{Y}))^2 \\
 &= E(\mathbf{x} \cdot \mathbf{Y})^2 - (\mathbf{x} \cdot E(\mathbf{Y}))^2 \\
 &= E(\mathbf{x} \cdot \mathbf{Y})^2 - (\mathbf{x} \cdot \boldsymbol{\tau})^2 \\
 &= E(\mathbf{Y}^\top \mathbf{A} \mathbf{Y}) - (\mathbf{x} \cdot \boldsymbol{\tau})^2 \\
 &= \text{tr}(\mathbf{A} \mathbf{I} \sigma^2) + \boldsymbol{\tau}^\top \mathbf{A} \boldsymbol{\tau} - (\mathbf{x} \cdot \boldsymbol{\tau})^2 \\
 &= \text{tr}(\mathbf{A} \sigma^2) + \underbrace{\boldsymbol{\tau}^\top \mathbf{A} \boldsymbol{\tau} - (\mathbf{x} \cdot \boldsymbol{\tau})^2}_{=0} \\
 &= x_1^2 \sigma^2 + x_2^2 \sigma^2 + \dots + x_N^2 \sigma^2 \\
 &= (x_1^2 + x_2^2 + \dots + x_N^2) \sigma^2 \\
 &= \|\mathbf{x}\|^2 \sigma^2.
 \end{aligned}$$

De onde segue o resultado. A seguir algumas observações.

Para verificar que a igualdade $\boldsymbol{\tau}^\top \mathbf{A} \boldsymbol{\tau} - (\mathbf{x} \cdot \boldsymbol{\tau})^2 = 0$ é verdadeira, vamos desenvolver para o caso $N = 2$, vejamos:

$$\begin{aligned}
 \boldsymbol{\tau}^\top \mathbf{A} \boldsymbol{\tau} - (\mathbf{x} \cdot \boldsymbol{\tau})^2 &= \boldsymbol{\tau}^\top \mathbf{A} \boldsymbol{\tau} - (\mathbf{x} \cdot \boldsymbol{\tau})(\mathbf{x} \cdot \boldsymbol{\tau}) \\
 &= \begin{pmatrix} \tau_1 & \tau_2 \end{pmatrix} \begin{pmatrix} x_1^2 & x_1 x_2 \\ x_2 x_1 & x_2^2 \end{pmatrix} \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} - \\
 &\quad \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \right] \left[\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \cdot \begin{pmatrix} \tau_1 \\ \tau_2 \end{pmatrix} \right] \\
 &= \tau_1^2 x_1^2 + 2\tau_1 \tau_2 x_1 x_2 + \tau_2^2 x_2^2 - (\tau_1^2 x_1^2 + 2\tau_1 \tau_2 x_1 x_2 + \tau_2^2 x_2^2) \\
 &= 0.
 \end{aligned}$$

ii) Usando o fato que $cov(X, Y) = E((X - E(X))(Y - E(Y)))$, obtemos:

$$\begin{aligned}
 cov(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) &= cov(\mathbf{x}^\top \mathbf{Y}, \mathbf{z}^\top \mathbf{Y}) \\
 &= E((\mathbf{x}^\top \mathbf{Y} - E(\mathbf{x}^\top \mathbf{Y}))(\mathbf{z}^\top \mathbf{Y} - E(\mathbf{z}^\top \mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) \mathbf{z}^\top (\mathbf{Y} - E(\mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top \mathbf{z}) \\
 &= \mathbf{x}^\top E((\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top) \mathbf{z} \\
 &= \mathbf{x}^\top cov(\mathbf{Y}) \mathbf{z} \\
 &= \mathbf{x}^\top \sigma^2 \mathbf{I} \mathbf{z} \\
 &= \mathbf{x}^\top (\sigma^2 \mathbf{z}) \\
 &= (\mathbf{x}^\top \mathbf{z}) \sigma^2 \\
 &= (\mathbf{x} \cdot \mathbf{z}) \sigma^2.
 \end{aligned}$$

iii) Primeiramente, vamos mostrar que $\mathbf{x} \cdot \mathbf{Y}$ é um estimador não viesado de $\mathbf{x} \cdot \boldsymbol{\tau}$,

$$E(\mathbf{x} \cdot \mathbf{Y}) = \mathbf{x} \cdot E(\mathbf{Y}) = \mathbf{x} \cdot \boldsymbol{\tau}.$$

O que mostra que esse estimador é não viesado.

Agora vamos mostrar que este é o melhor estimador. Suponha que $\mathbf{x}^\lambda \cdot \mathbf{Y}$ seja outro estimador não viesado de $\mathbf{x} \cdot \boldsymbol{\tau}$, com $\mathbf{x}^\lambda \in V_T$, sendo assim:

$$E(\mathbf{x}^\lambda \cdot \mathbf{Y}) = \mathbf{x}^\lambda \cdot E(\mathbf{Y}) = \mathbf{x}^\lambda \cdot \boldsymbol{\tau} \text{ e}$$

$$E(\mathbf{x} \cdot \mathbf{Y}) = \mathbf{x} \cdot E(\mathbf{Y}) = \mathbf{x} \cdot \boldsymbol{\tau}.$$

De onde vem que $\mathbf{x}^\lambda \cdot \boldsymbol{\tau} = \mathbf{x} \cdot \boldsymbol{\tau}$. Desse modo:

$$\mathbf{x}^\lambda \cdot \boldsymbol{\tau} - \mathbf{x} \cdot \boldsymbol{\tau} = 0$$

$$(\mathbf{x}^\lambda - \mathbf{x}) \cdot \boldsymbol{\tau} = 0.$$

Como $\boldsymbol{\tau} \in V_T$, temos que $\mathbf{x}^\lambda - \mathbf{x}$ deve ser ortogonal a um vetor genérico em V_T . Como \mathbf{x} e \mathbf{x}^λ também são vetores em V_T , $(\mathbf{x}^\lambda - \mathbf{x}) \cdot \boldsymbol{\tau} = 0$ só é possível se $\mathbf{x}^\lambda - \mathbf{x} = \mathbf{0}$. Portanto $\mathbf{x}^\lambda = \mathbf{x}$.

Resta mostrar agora que esse estimador é de variância mínima. Seja $\mathbf{z} \in V$ tal que $\mathbf{z} \cdot \mathbf{Y}$ é um estimador não viesado de $\mathbf{x} \cdot \boldsymbol{\tau}$, sendo assim:

$$E(\mathbf{z} \cdot \mathbf{Y}) = \mathbf{z} \cdot E(\mathbf{Y}) = \mathbf{z} \cdot \boldsymbol{\tau} \text{ e}$$

$$E(\mathbf{x} \cdot \mathbf{Y}) = \mathbf{x} \cdot E(\mathbf{Y}) = \mathbf{x} \cdot \boldsymbol{\tau}.$$

De onde vem que $\mathbf{z} \cdot \boldsymbol{\tau} = \mathbf{x} \cdot \boldsymbol{\tau}$. Desse modo:

$$\mathbf{z} \cdot \boldsymbol{\tau} - \mathbf{x} \cdot \boldsymbol{\tau} = 0$$

$$(\mathbf{z} - \mathbf{x}) \cdot \boldsymbol{\tau} = 0.$$

Como $\boldsymbol{\tau} \in V_T$, temos que $\mathbf{z} - \mathbf{x}$ deve ser ortogonal a um vetor genérico em V_T , então $\mathbf{z} - \mathbf{x}$ está no subespaço ortogonal a V_T . O subespaço ortogonal a V_T é seu complemento ortogonal V_T^\perp . Assim \mathbf{z} é da forma $\mathbf{z} = \mathbf{x} + (\mathbf{z} - \mathbf{x})$ com $\mathbf{x} \in V_T$ e $(\mathbf{z} - \mathbf{x}) \in V_T^\perp$, desse modo, por Pitágoras temos que:

$$\|\mathbf{z}\|^2 = \|\mathbf{x}\|^2 + \|\mathbf{z} - \mathbf{x}\|^2.$$

Por fim, calculando a variância dos estimadores, obtemos:

$$\text{var}(\mathbf{z} \cdot \mathbf{Y}) = \|\mathbf{z}\|^2 \sigma^2 \text{ e } \text{var}(\mathbf{x} \cdot \mathbf{Y}) = \|\mathbf{x}\|^2 \sigma^2.$$

Como $\|\mathbf{z}\|^2 \sigma^2 \geq \|\mathbf{x}\|^2 \sigma^2$ segue que $\mathbf{x} \cdot \mathbf{Y}$ possui variância mínima.

□

A covariância de duas variáveis aleatórias é uma medida de variabilidade conjunta dessas variáveis. Se $\mathbf{x} \cdot \mathbf{z} = 0$, ou seja, $\mathbf{x} \perp \mathbf{z}$, então $\text{cov}(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = 0$. Nesse caso as variáveis são não correlacionadas e as projeções de \mathbf{Y} em \mathbf{x} e \mathbf{z} são independentes.

Para se obter um estimador do vetor de parâmetros de tratamento $\boldsymbol{\tau}$ use o Teorema 3.1.3 item iii). Como em Bailey (2008), primeiro estime o parâmetro de tratamento τ_i . Para tanto, escreva $\mathbf{x} = \left(\frac{1}{r_i}\right) \mathbf{u}_i$. De onde vem que:

$$\mathbf{x} \cdot \boldsymbol{\tau} = \left(\frac{1}{r_i}\right) \mathbf{u}_i \cdot \boldsymbol{\tau} = \left(\frac{1}{r_i}\right) \text{SOMA}_{\tau=i} = \text{MÉDIA}_{\tau=i} = \tau_i.$$

Como:

$$\mathbf{x} \cdot \mathbf{Y} = \left(\frac{1}{r_i} \right) \mathbf{u}_i \cdot \mathbf{Y} = \left(\frac{1}{r_i} \right) \text{SOMA}_{T=i} = \text{MÉDIA}_{T=i},$$

temos que $\text{MÉDIA}_{T=i}$ é o melhor estimador linear não viesado de τ_i .

Assumindo que $\boldsymbol{\tau} \in V_T$, temos também que:

$$P_{V_T} \boldsymbol{\tau} = \boldsymbol{\tau} = \sum_{i=1}^t \tau_i \mathbf{u}_i.$$

Como vimos:

$$P_{V_T} \mathbf{Y} = \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i.$$

Portanto, $\sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i$ é o melhor estimador linear não viesado de $\boldsymbol{\tau}$.

Para se obter um estimador da variância, lembre-se de que a esperança do quadrado médio residual é dada pela Definição 2.4.2 como segue:

$$EQM(V_T^\perp) = \frac{E(\|P_{V_T^\perp} \mathbf{Y}\|^2)}{\dim(V_T^\perp)}.$$

Pelo Teorema 2.3.1, item ii), temos que:

$$EQM(V_T^\perp) = \frac{\|P_{V_T^\perp} \boldsymbol{\tau}\|^2 + (N-t)\sigma^2}{\dim V_T^\perp} = \frac{0 + (N-t)\sigma^2}{N-t} = \sigma^2.$$

Portanto, $QM(V_T^\perp)$ é um estimador não viesado de σ^2 .

No Teorema 3.1.3, item i), mostra-se que a variância do estimador $\mathbf{x} \cdot \mathbf{Y}$ é $\|\mathbf{x}\|^2 \sigma^2$. Seja $\mathbf{x} = \left(\frac{1}{r_i} \right) \mathbf{u}_i$, então:

$$\begin{aligned} \|\mathbf{x}\|^2 \sigma^2 &= \left\| \left(\frac{1}{r_i} \right) \mathbf{u}_i \right\|^2 \sigma^2 = \left(\frac{1}{r_i} \right) \mathbf{u}_i \cdot \left(\frac{1}{r_i} \right) \mathbf{u}_i \sigma^2 \\ &= \left(\frac{1}{r_i} \right)^2 \mathbf{u}_i \cdot \mathbf{u}_i \sigma^2 = \frac{1}{r_i^2} r_i \sigma^2 = \frac{1}{r_i} \sigma^2. \end{aligned}$$

Como $QM(V_T^\perp)$ é o estimador não viesado de σ^2 , segue que $\frac{1}{r_i} \sigma^2 = E \left(\frac{1}{r_i} QM(V_T^\perp) \right)$. Portanto, $\frac{1}{r_i} QM(V_T^\perp)$ é um estimador não viesado da variância do estimador $\mathbf{x} \cdot \mathbf{Y}$.

3.2 Delineamento em blocos casualizados

No delineamento em blocos casualizados, as parcelas do experimento não são homogêneas. Nesse sentido, é necessário agrupá-las em blocos os mais homogêneos possível, de modo que as parcelas dentro de cada bloco sejam aproximadamente iguais. Nesse caso, os tratamentos são alocados aleatoriamente às parcelas em cada bloco. De acordo com Bailey (2008), os blocos devem ter o mesmo tamanho e devem ser grandes o suficiente para que cada tratamento ocorra pelo menos uma vez em cada bloco.

Suponha que no delineamento em blocos, Ω seja composto de b blocos de tamanho k . Logo, $|\Omega| = N = b \times k$.

Exemplo

Veja um experimento projetado no delineamento em blocos casualizados, na Tabela 3.3. Nesse exemplo, existem $b = 4$ blocos. Note que os tratamentos A, B, C, D e E são alocados, aleatoriamente, aos blocos e os tratamentos ocorrem uma vez em cada bloco, assim $k = 5$. Portanto, $N = 4 \times 5 = 20$.

Tabela 3.3 – Exemplo experimento em blocos.

Bloco 1	D	A	B	C	E
Bloco 2	C	D	B	E	A
Bloco 3	B	A	E	D	C
Bloco 4	B	C	A	E	D

Fonte: Da autora (2019).

Para o delineamento em blocos o espaço vetorial V pode ser decomposto em quatro subespaços ortogonais. A análise de variância pode ser obtida, por meio das projeções ortogonais do vetor \mathbf{Y} nesses subespaços.

O subespaço V_B

Agora, além de V_T é possível definir o subespaço V_B . Também chamado de subespaço de bloco, V_B consiste de todos os vetores em V que são constantes em cada bloco. Para obter a projeção ortogonal de \mathbf{Y} em V_B é necessário obter uma base que gere esse subespaço. Tal base é obtida da seguinte maneira.

Para $j = 1, \dots, b$, seja \mathbf{v}_j o vetor cujo valor na parcela ω é igual a:

$$\begin{cases} 1, & \text{se } \omega \text{ está no bloco } j, \\ 0, & \text{caso contrário.} \end{cases}$$

Exemplo

Suponha um exemplo fictício em que existem três tratamentos A, B e C e dois blocos de tamanho quatro. As repetições são $r_A = 4$, $r_B = r_C = 2$ e $N = 8$. Na Tabela 3.4 mostra-se o conjunto das parcelas Ω , os blocos, o conjunto de tratamentos \mathfrak{T} , os vetores \mathbf{u}_A , \mathbf{u}_B e \mathbf{u}_C da base para V_T e os vetores \mathbf{v}_1 e \mathbf{v}_2 da base para V_B .

Tabela 3.4 – Vetores no experimento em blocos.

Ω	Blocos	\mathfrak{T}	\mathbf{u}_A	\mathbf{u}_B	\mathbf{u}_C	\mathbf{v}_1	\mathbf{v}_2
1	1	A	1	0	0	1	0
2	1	A	1	0	0	1	0
3	1	B	0	1	0	1	0
4	1	C	0	0	1	1	0
5	2	C	0	0	1	0	1
6	2	A	1	0	0	0	1
7	2	B	0	1	0	0	1
8	2	A	1	0	0	0	1

Fonte: Adaptado Bailey (2008).

Note que, $\mathbf{v}_j \cdot \mathbf{v}_j = k$ que é o tamanho do bloco, e $\mathbf{v}_j \cdot \mathbf{v}_l = 0$ para $j \neq l$. Sendo assim, $\{\mathbf{v}_1, \dots, \mathbf{v}_b\}$ é uma base para V_B e qualquer vetor \mathbf{v} em V_B é uma combinação linear única dos vetores dessa base. Logo, se existem b blocos $\dim(V_B) = b$.

Utilizando o Teorema 2.2.3 é possível obter a projeção de \mathbf{Y} em V_B . Para tanto, seja $\{\mathbf{v}_1, \dots, \mathbf{v}_b\}$ a base de V_B , então a projeção ortogonal de \mathbf{Y} em V_B é dada por:

$$\begin{aligned}
 P_{V_B}\mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1}\right) \mathbf{v}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{v}_b}{\mathbf{v}_b \cdot \mathbf{v}_b}\right) \mathbf{v}_b \\
 &= \sum_{j=1}^b \left(\frac{\mathbf{Y} \cdot \mathbf{v}_j}{\mathbf{v}_j \cdot \mathbf{v}_j}\right) \mathbf{v}_j \\
 &= \sum_{j=1}^b \left(\frac{\text{SOMA}_{B=j}}{k}\right) \mathbf{v}_j \\
 &= \sum_{j=1}^b (\text{MÉDIA}_{B=j}) \mathbf{v}_j \\
 &= \text{vetor de médias dos blocos.}
 \end{aligned}$$

O subespaço W_B

Como $V_0 \subset V_B$ é possível definir o subespaço W_B , também chamado de subespaço de efeito de bloco, dado pelo seguinte conjunto de vetores:

$$W_B = V_B \cap V_0^\perp = \{\mathbf{v} \in V_B : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

Note que o subespaço W_B é constituído, por todos os vetores em V_B que são ortogonais a V_0 . O subespaço V_B fica decomposto, então, na soma direta de V_0 e W_B , ou seja:

$$V_B = V_0 \oplus W_B.$$

Isso implica em:

$$\begin{aligned}
 \dim(W_B) &= \dim(V_B) - \dim(V_0) \\
 &= b - 1.
 \end{aligned}$$

A decomposição de V_B leva também à decomposição de vetores em V_B . Desse modo, a projeção ortogonal de \mathbf{Y} em W_B pode ser obtida pela diferença entre $P_{V_B}\mathbf{Y}$ e $P_{V_0}\mathbf{Y}$, como segue:

$$\begin{aligned} P_{W_B}\mathbf{Y} &= P_{V_B}\mathbf{Y} - P_{V_0}\mathbf{Y} \\ &= \sum_{j=1}^b (\text{MÉDIA}_{B=j}) \mathbf{v}_j - \bar{Y}\mathbf{u}_0 \\ &= \text{vetor de médias dos blocos} - \text{vetor de média geral} \\ &= \text{vetor de efeitos dos blocos.} \end{aligned}$$

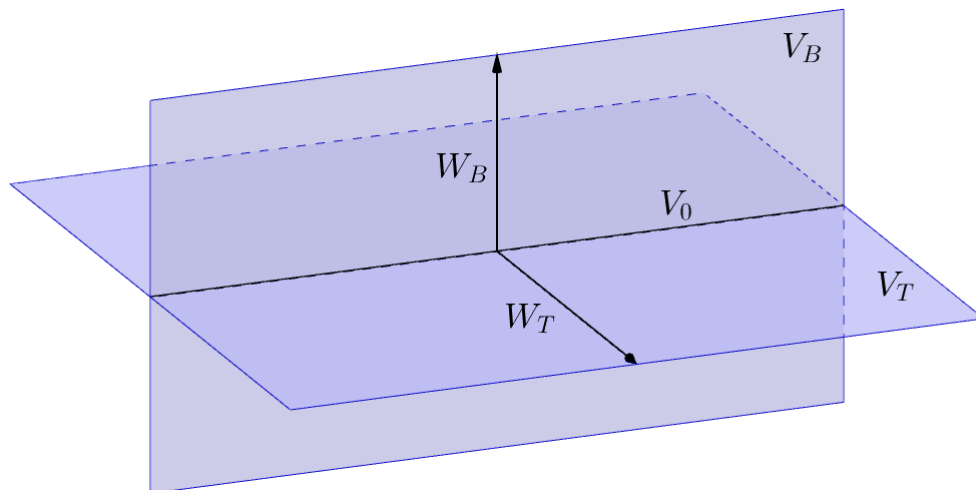
O subespaço W_E

Antes de definir o subespaço W_E é necessário trazer algumas definições e teoremas.

Definição 3.2.1. *Um delineamento em blocos é dito ortogonal se os subespaços W_T e W_B forem ortogonais entre si.*

Veja os subespaços ortogonais W_T e W_B na Figura 3.9.

Figura 3.9 – Subespaços ortogonais para delineamento em blocos.



Fonte: Da autora (2019).

Teorema 3.2.1. *Dado um delineamento em blocos para t tratamentos em b blocos de tamanho k , seja s_{ij} o número de vezes que o tratamento i ocorre no bloco j , para $i = 1, \dots, t$ e $j = 1, \dots, b$. Então, o delineamento em blocos é ortogonal se, e somente se, $s_{ij} = \frac{r_i}{b}$ para $i = 1, \dots, t$ e $j = 1, \dots, b$.*

A demonstração desse teorema envolve fundamentos combinatórios e encontra-se em Bailey (2008).

Definição 3.2.2. *Um delineamento é dito em blocos completos se possui blocos de tamanho t , com cada tratamento ocorrendo uma única vez em cada bloco.*

Corolário 3.2.1. *Os delineamentos em blocos completos são ortogonais.*

Para o delineamento de blocos ortogonais, W_T e W_B são ortogonais entre si. Nesse caso, é possível construir o subespaço W_E , também chamado de subespaço de resíduo.

Primeiro, lembre-se da definição da soma de dois subespaços, que é dada por:

$$V_T + V_B = \{\mathbf{w}_1 + \mathbf{w}_2 : \mathbf{w}_1 \in V_T \text{ e } \mathbf{w}_2 \in V_B\}.$$

Como $V_T = V_0 \oplus W_T$ e $V_B = V_0 \oplus W_B$, temos que:

$$V_T + V_B = V_0 \oplus W_T \oplus W_B.$$

Agora é possível definir o subespaço W_E , que consiste no conjunto de vetores:

$$W_E = \{\mathbf{v} \in V : \mathbf{v} \text{ é ortogonal a } \mathbf{w} \text{ para todo } \mathbf{w} \in V_T + V_B\} = (V_T + V_B)^\perp.$$

Note que W_E é constituído pelos vetores em V que são ortogonais a $V_T + V_B$. Definidos esses subespaços, vimos que é possível decompor V como soma direta de quatro subespaços, os quais são:

$$\begin{aligned} V &= (V_T + V_B) \oplus (V_T + V_B)^\perp \\ &= V_0 \oplus W_T \oplus W_B \oplus W_E. \end{aligned}$$

Assim, a dimensão de V corresponde à soma das dimensões dos quatro subespaços. De modo que a dimensão de W_E pode ser obtida por diferença:

$$\begin{aligned} \dim(W_E) &= \dim(V) - \dim(V_0) - \dim(W_T) - \dim(W_B) \\ &= N - 1 - (t - 1) - (b - 1) \\ &= N - t - b + 1. \end{aligned}$$

A decomposição de V é única, o que significa que todo vetor em V se escreve de modo único como soma de vetores desses subespaços. Desse modo, o vetor \mathbf{Y} pode ser decomposto em soma de quatro vetores de projeção:

$$\mathbf{Y} = P_{V_0}\mathbf{Y} + P_{W_T}\mathbf{Y} + P_{W_B}\mathbf{Y} + P_{W_E}\mathbf{Y}.$$

Portanto, o vetor de resíduos é dado por:

$$\begin{aligned} P_{W_E}\mathbf{Y} &= \mathbf{Y} - P_{V_0}\mathbf{Y} - P_{W_T}\mathbf{Y} - P_{W_B}\mathbf{Y} \\ &= \mathbf{Y} - \bar{Y}\mathbf{u}_0 - \left(\sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \bar{Y}\mathbf{u}_0 \right) - \left(\sum_{j=1}^b (\text{MÉDIA}_{B=j}) \mathbf{v}_j - \bar{Y}\mathbf{u}_0 \right) \\ &= \mathbf{Y} - \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \sum_{j=1}^b (\text{MÉDIA}_{B=j}) \mathbf{v}_j + \bar{Y}\mathbf{u}_0 \\ &= \text{vetor de dados} - \text{vetor de médias dos tratamentos} \\ &\quad - \text{vetor de médias dos blocos} + \text{vetor de média geral.} \end{aligned}$$

Novamente, é possível usar a Definição 2.4.1, para obter as somas de quadrados, os graus de liberdade e os quadrados médios para os subespaços definidos anteriormente. Para isso, basta se obter as normas ao quadrado dos vetores de projeção e dividir pela dimensão dos respectivos subespaços.

As soma de quadrados e os quadrados médios para V_0 e W_T são definidas de modo análogo à Seção 3.1.

A soma de quadrados para W_B , também chamada de soma de quadrados dos efeitos dos blocos pode ser obtida por diferença, pois $P_{V_B}\mathbf{Y}$ é soma de vetores ortogonais $P_{V_0}\mathbf{Y}$ e $P_{W_B}\mathbf{Y}$ e por Pitágoras temos que:

$$\begin{aligned} \|P_{V_B}\mathbf{Y}\|^2 &= \|P_{V_0}\mathbf{Y}\|^2 + \|P_{W_B}\mathbf{Y}\|^2 \\ \|P_{W_B}\mathbf{Y}\|^2 &= \|P_{V_B}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2, \end{aligned}$$

em que $\|P_{V_B}\mathbf{Y}\|^2$ é a soma de quadrados para V_B , também chamada de soma de quadrados das médias dos blocos, dada por:

$$\begin{aligned}
 \|P_{V_B}\mathbf{Y}\|^2 &= P_{V_B}\mathbf{Y} \cdot P_{V_B}\mathbf{Y} \\
 &= \sum_{j=1}^b \left(\frac{\text{SOMA}_{B=j}}{k} \right) \mathbf{v}_j \cdot \sum_{j=1}^b \left(\frac{\text{SOMA}_{B=j}}{k} \right) \mathbf{v}_j \\
 &= \sum_{j=1}^b \left(\frac{\text{SOMA}_{B=j}}{k} \right)^2 \mathbf{v}_j \cdot \mathbf{v}_j \\
 &= \sum_{j=1}^b \frac{(\text{SOMA}_{B=j})^2}{k^2} k \\
 &= \sum_{j=1}^b \frac{(\text{SOMA}_{B=j})^2}{k}.
 \end{aligned}$$

O quadrado médio W_B é dado por:

$$QM(W_B) = \frac{\|P_{W_B}\mathbf{Y}\|^2}{\dim(W_B)} = \frac{\|P_{W_B}\mathbf{Y}\|^2}{b-1}.$$

A soma de quadrados para W_E , que é a soma de quadrados dos resíduos, pode ser obtida por diferença, isso porque \mathbf{Y} é soma dos vetores $P_{V_0}\mathbf{Y}$, $P_{W_T}\mathbf{Y}$, $P_{W_B}\mathbf{Y}$ e $P_{W_E}\mathbf{Y}$. Desse modo, temos que:

$$\|\mathbf{Y}\|^2 = \|P_{V_0}\mathbf{Y}\|^2 + \|P_{W_T}\mathbf{Y}\|^2 + \|P_{W_B}\mathbf{Y}\|^2 + \|P_{W_E}\mathbf{Y}\|^2.$$

Logo, a soma de quadrados para W_E é dada por:

$$\begin{aligned}
 \|P_{W_E}\mathbf{Y}\|^2 &= \|\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 - \|P_{W_T}\mathbf{Y}\|^2 - \|P_{W_B}\mathbf{Y}\|^2 \\
 &= \|\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 - (\|P_{W_T}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2) - (\|P_{W_B}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2) \\
 &= \|\mathbf{Y}\|^2 - \|P_{V_T}\mathbf{Y}\|^2 - \|P_{V_B}\mathbf{Y}\|^2 + \|P_{V_0}\mathbf{Y}\|^2.
 \end{aligned}$$

O quadrado médio W_E é dado por:

$$QM(W_E) = \frac{\|P_{W_E}\mathbf{Y}\|^2}{\dim(W_E)} = \frac{\|P_{W_E}\mathbf{Y}\|^2}{N-t-b+1}.$$

Realizados esses cálculos, é possível se obter a análise de variância para o delineamento em blocos casualizados, conforme a Tabela 3.5. Note que as fontes de variação da tabela são os subespaços da decomposição V como soma direta.

Tabela 3.5 – Tabela de análise de variância para delineamento em blocos casualizados.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	$QM(V_0)/QM(W_E)$
Tratamentos W_T	$t - 1$	$\ P_{W_T}\mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/QM(W_E)$
Blocos W_B	$b - 1$	$\ P_{W_B}\mathbf{Y}\ ^2$	$QM(W_B)$	$QM(W_B)/QM(W_E)$
Resíduo W_E	$N - t - b + 1$	$\ P_{W_E}\mathbf{Y}\ ^2$	$QM(W_E)$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

Assuma que $\mathbf{Y} \sim N(\boldsymbol{\tau} + \boldsymbol{\zeta}, \sigma^2\mathbf{I})$ onde $\boldsymbol{\tau} \in V_T$ e $\boldsymbol{\zeta} \in V_B$, de modo que o modelo aditivo para soma de vetores é correto e que qualquer variação no experimento é decorrente do erro. Em termos vetoriais temos:

$$\mathbf{Y} = \boldsymbol{\tau} + \boldsymbol{\zeta} + \boldsymbol{\varepsilon},$$

em que $E(\mathbf{Y}) = \boldsymbol{\tau} + \boldsymbol{\zeta} \in V_T + V_B$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ e $cov(\mathbf{Y}) = \sigma^2\mathbf{I}$.

Note que agora:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_T \text{ onde } \boldsymbol{\tau}_0 \in V_0 \text{ e } \boldsymbol{\tau}_T \in W_T \text{ e}$$

$$\boldsymbol{\zeta} = \boldsymbol{\zeta}_0 + \boldsymbol{\zeta}_B \text{ onde } \boldsymbol{\zeta}_0 \in V_0 \text{ e } \boldsymbol{\zeta}_B \in W_B.$$

Assim sendo, $E(\mathbf{Y}) = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_T + \boldsymbol{\zeta}_0 + \boldsymbol{\zeta}_B$, e como $\boldsymbol{\tau}_0$ e $\boldsymbol{\zeta}_0$ estão em V_0 , ambos são múltiplos de \boldsymbol{u}_0 , de modo que não podem ser distinguidos nos dados. Como os efeitos de tratamentos e de blocos, encontram-se em W_T e W_B respectivamente, e esses são ortogonais, é possível estimar combinações lineares dos efeitos, uma vez que:

$$\boldsymbol{\tau}_T = \boldsymbol{\tau} - \boldsymbol{\tau}_0 \text{ e } \boldsymbol{\zeta}_B = \boldsymbol{\zeta} - \boldsymbol{\zeta}_0.$$

O teste de hipótese

Agora temos três razões de variâncias entre quadrados médios para realizar o teste F . Como já vimos, o teste para verificar se a média geral é zero não tem interesse prático.

Para testar se existe diferença entre as médias dos tratamentos, pode-se testar a hipótese nula H_0 , contra a hipótese alternativa H_a , as quais são:

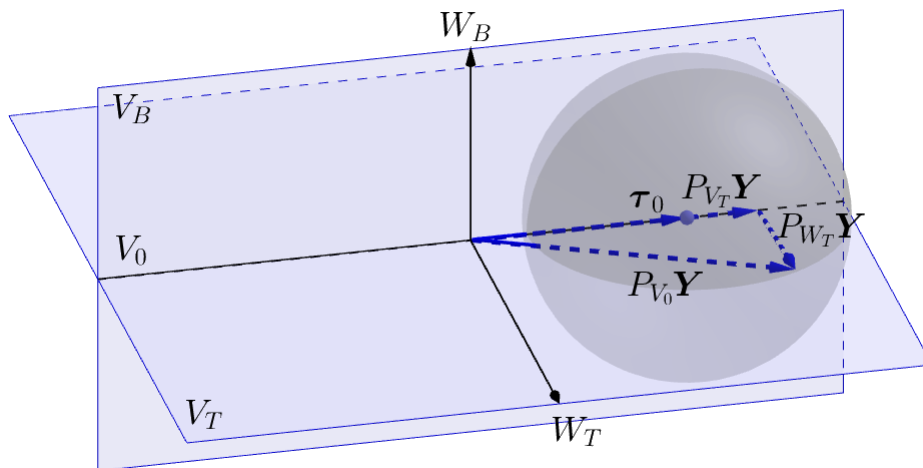
$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_a : \boldsymbol{\tau} \text{ não é um vetor constante.}$$

Se a hipótese nula é verdadeira $\boldsymbol{\tau}$ é um vetor constante, logo, $\boldsymbol{\tau} \in V_0$, e portanto $\boldsymbol{\tau} = \boldsymbol{\tau}_0$. Se a hipótese nula é falsa $\boldsymbol{\tau}$ não é um vetor constante, logo, $\boldsymbol{\tau} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_T$, e portanto, existe efeito de tratamento. Sob H_0 verdadeira, a nuvem de pontos é centrada em $\boldsymbol{\tau}_0$ que pertence a V_0 , como mostra a Figura 3.10.

Para realizar o teste, compare o vetor $P_{W_T} \mathbf{Y}$ com o vetor $P_{W_E} \mathbf{Y}$ (Figura 3.10). Para tanto, compare a razão entre $QM(W_T)$ e $QM(W_E)$. Se esse valor for menor que um, não se rejeita H_0 e pode-se concluir que não existe diferença entre as médias dos tratamentos. Se esse valor for maior que um, rejeita-se H_0 e pode-se concluir que existe diferença entre as médias dos tratamentos.

Figura 3.10 – Teste de hipótese para tratamentos.



Fonte: Da autora (2019).

O teste usado para verificar se existe diferença entre as médias dos blocos, na maioria das vezes, não é de interesse do pesquisador. Isso porque, ao planejar o delineamento em blocos já se sabe que não existe homogeneidade entre os blocos, então, certamente existe diferença entre as médias dos blocos. Entretanto, caso se pretenda fazer o teste, pode-se testar a hipótese nula H_0 , contra a hipótese alternativa H_a , as quais são:

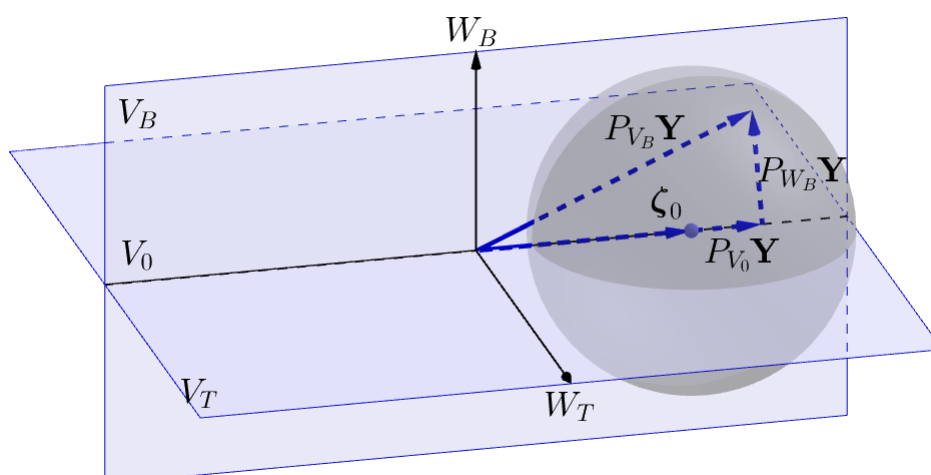
$$H_0 : \zeta_1 = \zeta_2 = \dots = \zeta_b$$

H_a : ζ não é um vetor constante.

Se a hipótese nula é verdadeira ζ é um vetor constante, logo, $\zeta \in V_0$ e, portanto $\zeta = \zeta_0$. Se a hipótese nula é falsa ζ não é um vetor constante, logo, $\zeta = \zeta_0 + \zeta_B$ e, portanto, existe efeito de bloco. Sob H_0 verdadeira, a nuvem de pontos é centrada em ζ_0 que pertence a V_0 , como mostra a Figura 3.11.

Para realizar o teste, compare o vetor $P_{W_B}\mathbf{Y}$ com o vetor $P_{W_E}\mathbf{Y}$ (Figura 3.11). Para tanto, compare a razão entre $QM(W_B)$ e $QM(W_E)$. Se esse valor for menor que um, não se rejeita H_0 e pode-se concluir que não existe diferença entre as médias dos blocos. Se esse valor for maior que um, rejeita-se H_0 e pode-se concluir que existe diferença entre as médias dos blocos.

Figura 3.11 – Teste de hipótese para blocos.



Fonte: Da autora (2019).

3.3 Delineamento em linhas e colunas

No delineamento em linhas e colunas, as parcelas do experimento possuem duas fontes não homogêneas, de modo que, é necessário controlar duas fontes de variação no experimento. Nesse sentido, esse delineamento utiliza controle local duplo e os tratamentos são alocados, aleatoriamente, nas linhas e nas colunas.

Assim como em Bailey (2008), assumo que no delineamento em linhas e colunas, cada linha encontra cada coluna em uma única parcela, todos os tratamentos ocorrem com igual frequência em cada linha e em cada coluna, existem m linhas e n colunas. Logo, o número total de parcelas é $N = m \times n$. Note que o delineamento é projetado em um retângulo.

Exemplo

Veja o um experimento projetado no delineamento em linhas e colunas na Tabela 3.6. Nesse exemplo, $m = 4$, $n = 8$ e, portanto $N = m \times n = 4 \times 8 = 32$. Os tratamentos A, B, C e D são alocados, aleatoriamente, às linhas e às colunas, de modo que, os tratamentos repetem nas linhas e não repetem nas colunas. Assim, os tratamentos têm repetição $r = m \times n / t = 32 / 4 = 8$.

Tabela 3.6 – Exemplo experimento linhas e colunas.

	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5	Coluna 6	Coluna 7	Coluna 8
Linha 1	D	A	B	C	D	A	B	C
Linha 2	B	C	D	A	B	C	D	A
Linha 3	A	B	C	D	A	B	C	D
Linha 4	C	D	A	B	C	D	A	B

Fonte: Da autora (2019).

Quadrado Latino

O quadrado latino é um caso particular do delineamento em linhas e colunas, que ocorre quando cada tratamento acontece uma vez em cada linha e uma vez em cada coluna. Nesse caso $n = m = t$. De acordo com Pimentel e Garcia (2002), no quadrado latino, cada linha e cada coluna é um bloco completo. Quando planejado, ele controla heterogeneidade em duas direções perpendiculares entre si.

Exemplo

Veja o experimento projetado em quadrado latino 5×5 na Tabela 3.7. Note que, nesse exemplo, $m = n = t = 5$. Os tratamentos A, B, C, D, e E são alocados aleatoriamente às linhas e às colunas, de modo que, os tratamentos não se repetem nas linhas e nem nas colunas.

Tabela 3.7 – Exemplo experimento em quadrado latino.

	Coluna 1	Coluna 2	Coluna 3	Coluna 4	Coluna 5
Linha 1	B	E	D	A	C
Linha 2	C	A	B	D	E
Linha 3	D	B	C	E	A
Linha 4	A	C	E	B	D
Linha 5	E	D	A	C	B

Fonte: Pimentel-Gomes e Garcia (2002).

Perceba que a aleatorização dos tratamentos nas parcelas não é trivial. De modo que é necessário mostrar métodos de construção para delineamentos em linhas e colunas.

Para o quadrado latino, a construção mais comum é pelo método cíclico. Para a construção por esse método, primeiro colocam-se os tratamentos na linha superior, em qualquer ordem. Em seguida, na segunda linha, desloque todos os tratamentos uma parcela para a direita, continua-se deslocando os tratamentos até a última linha. Veja o exemplo de quadrado latino cíclico de ordem 4 na Figura 3.12.

Figura 3.12 – Construção quadrado latino método cíclico.

A	B	C	D
D	A	B	C
C	D	A	B
B	C	D	A

Fonte: Bailey (2008).

Para o caso geral, em que os números de linhas e de colunas são diferentes, como aponta Bailey (2008), a construção do delineamento em linhas e colunas pode ser feita da seguinte maneira:

- i) Divida o retângulo $m \times n$ em quadrados $t \times t$;
- ii) Em cada quadrado $t \times t$ coloque um quadrado latino de ordem t (use os mesmos símbolos em cada quadrado);
- iii) Aleatoriamente permute as m linhas (não os tratamentos dentro delas);
- iv) Aleatoriamente permute as n colunas (não os tratamentos dentro delas).

Veja o exemplo da Figura 3.13, para o delineamento em linhas e colunas 3×6 com $t = 3$.

Subespaços ortogonais

Novamente, é possível definir os subespaços ortogonais para decomposição de V para o delineamento em linhas e colunas. Os subespaços V_T , W_T e V_0 são definidos de forma análoga ao que foi feito na Seção 3.1 e possuem dimensão t , $t - 1$ e 1 respectivamente. As projeções de \mathbf{Y} em V_T , W_T e V_0 também são análogas ao que foi feito na Seção 3.1.

Assim como foi definido o subespaço de blocos na Seção 3.2, é possível definir os subespaços de linhas e de colunas.

Os subespaços V_L e V_C

Figura 3.13 – Construção do delineamento em linhas e colunas para o caso geral.

A	B	C	A	B	C
C	A	B	C	A	B
B	C	A	B	C	A

permutação das linhas

↓

B	C	A	B	C	A
A	B	C	A	B	C
C	A	B	C	A	B

permutação das colunas

↓

C	C	A	B	B	A
B	B	C	A	A	C
A	A	B	C	C	B

Fonte: Da autora (2019).

O subespaço V_L , também chamado de subespaço de linha, consiste naqueles vetores em V que são constantes em cada linha. Note que se existem m linhas, temos que $\dim(V_L) = m$.

O subespaço V_C , também chamado de subespaço de coluna, consiste naqueles vetores em V que são constantes em cada coluna. Note que se existem n colunas, temos que $\dim(V_C) = n$.

Exemplo

Suponha um exemplo fictício em que existem dois tratamentos A e B, duas linhas e quatro colunas. Desse modo, o delineamento em linhas e colunas é 2×4 e as repetições são $r_A = 4$ e $r_B = 4$. O experimento foi aleatorizado conforme a Figura 3.14. Na Tabela 3.8, mostra-se o conjunto de parcelas Ω , as linhas, as colunas, o conjunto de tratamentos \mathfrak{T} , os vetores \mathbf{u}_A e \mathbf{u}_B da base para V_T , os vetores \mathbf{l}_1 e \mathbf{l}_2 da base para V_L e os vetores \mathbf{c}_1 , \mathbf{c}_2 , \mathbf{c}_3 e \mathbf{c}_4 da base para V_C .

Figura 3.14 – Exemplo aleatorizado linhas e colunas.

B	B	A	A
A	A	B	B

Fonte: Da autora (2019).

Tabela 3.8 – Vetores no experimento em linhas e colunas.

Ω	Linhas	Colunas	\mathfrak{T}	\mathbf{u}_A	\mathbf{u}_B	\mathbf{l}_1	\mathbf{l}_2	\mathbf{c}_1	\mathbf{c}_2	\mathbf{c}_3	\mathbf{c}_4
1	1	1	B	0	1	1	0	1	0	0	0
2	1	2	B	0	1	1	0	0	1	0	0
3	1	3	A	1	0	1	0	0	0	1	0
4	1	4	A	1	0	1	0	0	0	0	1
5	2	1	A	1	0	0	1	1	0	0	0
6	2	2	A	1	0	0	1	0	1	0	0
7	2	3	B	0	1	0	1	0	0	1	0
8	2	4	B	0	1	0	1	0	0	0	1

Fonte: Da autora (2019).

Nota-se que, para um vetor \mathbf{l}_j da base de V_L , temos que $\mathbf{l}_j \cdot \mathbf{l}_j = n$, que é o número de colunas. Analogamente, para um vetor \mathbf{c}_k da base de V_C , temos que $\mathbf{c}_k \cdot \mathbf{c}_k = m$, que é o número de linhas.

Agora é possível usar o Teorema 2.2.3 para obter as projeções de \mathbf{Y} em V_L e V_C .

Seja $\{\mathbf{l}_1, \dots, \mathbf{l}_m\}$ uma base para V_L , a projeção ortogonal de \mathbf{Y} em V_L é dada por:

$$\begin{aligned}
 P_{V_L} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{l}_1}{\mathbf{l}_1 \cdot \mathbf{l}_1} \right) \mathbf{l}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{l}_m}{\mathbf{l}_m \cdot \mathbf{l}_m} \right) \mathbf{l}_m \\
 &= \sum_{j=1}^m \left(\frac{\mathbf{Y} \cdot \mathbf{l}_j}{\mathbf{l}_j \cdot \mathbf{l}_j} \right) \mathbf{l}_j \\
 &= \sum_{j=1}^m \left(\frac{\text{SOMA}_{L=j}}{n} \right) \mathbf{l}_j \\
 &= \sum_{j=1}^m (\text{MÉDIA}_{L=j}) \mathbf{l}_j \\
 &= \text{vetor de médias das linhas.}
 \end{aligned}$$

Seja $\{\mathbf{c}_1, \dots, \mathbf{c}_n\}$ uma base para V_C , a projeção ortogonal de \mathbf{Y} em V_C é dada por:

$$\begin{aligned}
 P_{V_C}\mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{c}_1}{\mathbf{c}_1 \cdot \mathbf{c}_1}\right) \mathbf{c}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{c}_n}{\mathbf{c}_n \cdot \mathbf{c}_n}\right) \mathbf{c}_n \\
 &= \sum_{k=1}^n \left(\frac{\mathbf{Y} \cdot \mathbf{c}_k}{\mathbf{c}_k \cdot \mathbf{c}_k}\right) \mathbf{c}_k \\
 &= \sum_{k=1}^n \left(\frac{\text{SOMA}_{C=k}}{m}\right) \mathbf{c}_k \\
 &= \sum_{k=1}^n (\text{MÉDIA}_{C=k}) \mathbf{c}_k \\
 &= \text{vetor de médias das colunas.}
 \end{aligned}$$

Os subespaços W_L e W_C

A partir de V_L é possível definir o subespaço W_L , também chamado de subespaço de efeito de linha, como segue:

$$W_L = V_L \cap V_0^\perp = \{\mathbf{v} \in V_L : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

Note que W_L é constituído pelos vetores em V_L que são perpendiculares a V_0 . Agora o subespaço V_L é decomposto como soma direta de V_0 e W_L , ou seja:

$$V_L = V_0 \oplus W_L.$$

Então, a dimensão de W_L pode ser obtida por diferença, como segue:

$$\begin{aligned}
 \dim(W_L) &= \dim(V_L) - \dim(V_0) \\
 &= m - 1.
 \end{aligned}$$

A partir de V_C é possível definir o subespaço W_C , também chamado de subespaço de efeito de coluna, como segue:

$$W_C = V_C \cap V_0^\perp = \{\mathbf{v} \in V_C : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

Note que W_C é constituído pelos vetores em V_C que são perpendiculares a V_0 . Agora o subespaço V_C é decomposto como soma direta de V_0 e W_C , ou seja:

$$V_C = V_0 \oplus W_C.$$

Então a dimensão de W_C pode ser obtida por diferença, como segue:

$$\begin{aligned} \dim(W_C) &= \dim(V_C) - \dim(V_0) \\ &= n - 1. \end{aligned}$$

As decomposições de V_L e V_C levam a decomposições de vetores nesses subespaços. Desse modo, os vetores de projeção de \mathbf{Y} em W_L e W_C podem ser obtidos por diferença.

$$\begin{aligned} P_{W_L} \mathbf{Y} &= P_{V_L} \mathbf{Y} - P_{V_0} \mathbf{Y} \\ &= \sum_{j=1}^m (\text{MÉDIA}_{L=j}) \mathbf{l}_j - \bar{Y} \mathbf{u}_0 \\ &= \text{vetor de médias das linhas} - \text{vetor de média geral} \\ &= \text{vetor de efeitos das linhas.} \end{aligned}$$

$$\begin{aligned} P_{W_C} \mathbf{Y} &= P_{V_C} \mathbf{Y} - P_{V_0} \mathbf{Y} \\ &= \sum_{k=1}^n (\text{MÉDIA}_{C=k}) \mathbf{c}_k - \bar{Y} \mathbf{u}_0 \\ &= \text{vetor de médias das colunas} - \text{vetor de média geral} \\ &= \text{vetor de efeitos das colunas.} \end{aligned}$$

O subespaço W_E

Se cada linha encontra cada coluna em uma única parcela e todos os tratamentos ocorrem com igual frequência em cada linha e em cada coluna. Nessas condições, pelo Teorema 3.2.1, os subespaços W_T , W_L e W_C são ortogonais entre si.

Conforme na Seção 3.2, é possível definir a soma dos subespaços V_T , V_L e V_C . Lembre-se que $V_T = V_0 \oplus W_T$, $V_L = V_0 \oplus W_L$ e $V_C = V_0 \oplus W_C$, o que resulta em:

$$V_T + V_L + V_C = V_0 \oplus W_T \oplus W_L \oplus W_C.$$

Agora, é possível definir o subespaço W_E , também chamado de subespaço de resíduo, que consiste no seguinte conjunto de vetores:

$$\begin{aligned} W_E &= \{ \mathbf{v} \in V : \mathbf{v} \text{ é ortogonal a } \mathbf{w} \text{ para todo } \mathbf{w} \in V_T + V_L + V_C \} \\ &= (V_T + V_L + V_C)^\perp. \end{aligned}$$

Definidos esses subespaços, vimos que é possível decompor V como soma direta de cinco subespaços, os quais são:

$$\begin{aligned} V &= (V_T + V_L + V_C) \oplus (V_T + V_L + V_C)^\perp \\ &= V_0 \oplus W_T \oplus W_L \oplus W_C \oplus W_E. \end{aligned}$$

Desse modo, a dimensão de V corresponde à soma das dimensões desses cinco subespaços. Portanto, a dimensão de W_E pode ser obtida por diferença, como a seguir:

$$\begin{aligned} \dim(W_E) &= \dim(V) - \dim(V_0) - \dim(W_T) - \dim(W_L) - \dim(W_C) \\ &= N - 1 - (t - 1) - (m - 1) - (n - 1) \\ &= N - t - m - n + 2. \end{aligned}$$

A partir da decomposição de V o vetor \mathbf{Y} pode ser decomposto como soma dos vetores:

$$\mathbf{Y} = P_{V_0}\mathbf{Y} + P_{W_T}\mathbf{Y} + P_{W_L}\mathbf{Y} + P_{W_C}\mathbf{Y} + P_{W_E}\mathbf{Y}.$$

Logo, o vetor de resíduos $P_{W_E} \mathbf{Y}$ pode ser obtido por diferença:

$$\begin{aligned}
 P_{W_E} \mathbf{Y} &= \mathbf{Y} - P_{V_0} \mathbf{Y} - P_{W_T} \mathbf{Y} - P_{W_L} \mathbf{Y} - P_{W_C} \mathbf{Y} \\
 &= \mathbf{Y} - \bar{Y} \mathbf{u}_0 - \left(\sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \bar{Y} \mathbf{u}_0 \right) \\
 &\quad - \left(\sum_{j=1}^m (\text{MÉDIA}_{L=j}) \mathbf{l}_j - \bar{Y} \mathbf{u}_0 \right) - \left(\sum_{k=1}^n (\text{MÉDIA}_{C=k}) \mathbf{c}_k - \bar{Y} \mathbf{u}_0 \right) \\
 &= \mathbf{Y} - \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \sum_{j=1}^m (\text{MÉDIA}_{L=j}) \mathbf{l}_j - \sum_{k=1}^n (\text{MÉDIA}_{C=k}) \mathbf{c}_k + 2(\bar{Y} \mathbf{u}_0) \\
 &= \text{vetor de dados} - \text{vetor de médias dos tratamentos} - \text{vetor de médias das linhas} \\
 &\quad - \text{vetor de médias das colunas} - 2(\text{vetor de média geral}).
 \end{aligned}$$

Agora, use a Definição 2.4.1 para obter as somas de quadrados, os graus de liberdade e os quadrados médios para as fontes de variação da análise de variância.

As somas de quadrados e os quadrados médios para V_0 e W_T são análogas ao que foi feito na Seção 3.1.

A soma de quadrados para W_L , ou soma de quadrados dos efeitos das linhas, é dada por:

$$\|P_{W_L} \mathbf{Y}\|^2 = \|P_{V_L} \mathbf{Y}\|^2 - \|P_{V_0} \mathbf{Y}\|^2,$$

em $\|P_{V_L} \mathbf{Y}\|^2$ é a soma de quadrados para V_L , ou soma de quadrados das médias das linhas, dada por:

$$\begin{aligned}
 \|P_{V_L} \mathbf{Y}\|^2 &= P_{V_L} \mathbf{Y} \cdot P_{V_L} \mathbf{Y} \\
 &= \sum_{j=1}^m \left(\frac{\text{SOMA}_{L=j}}{n} \right) \mathbf{l}_j \cdot \sum_{j=1}^m \left(\frac{\text{SOMA}_{L=j}}{n} \right) \mathbf{l}_j \\
 &= \sum_{j=1}^m \left(\frac{\text{SOMA}_{L=j}}{n} \right)^2 \mathbf{l}_j \cdot \mathbf{l}_j \\
 &= \sum_{j=1}^m \frac{(\text{SOMA}_{L=j})^2}{n^2} n \\
 &= \sum_{j=1}^m \frac{(\text{SOMA}_{L=j})^2}{n}.
 \end{aligned}$$

Assim, o quadrado médio para W_L é dado por:

$$QM(W_L) = \frac{\|P_{W_L}\mathbf{Y}\|^2}{\dim(W_L)} = \frac{\|P_{W_L}\mathbf{Y}\|^2}{m-1}.$$

A soma de quadrados para W_C , ou soma de quadrados dos efeitos das colunas, é dada por:

$$\|P_{W_C}\mathbf{Y}\|^2 = \|P_{V_C}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2,$$

em que $\|P_{V_C}\mathbf{Y}\|^2$ é a soma de quadrados para V_C , ou soma de quadrados das médias das colunas, dada por:

$$\begin{aligned} \|P_{V_C}\mathbf{Y}\|^2 &= P_{V_C}\mathbf{Y} \cdot P_{V_C}\mathbf{Y} \\ &= \sum_{k=1}^n \left(\frac{\text{SOMA}_{C=k}}{m} \right) \mathbf{c}_k \cdot \sum_{k=1}^n \left(\frac{\text{SOMA}_{C=k}}{m} \right) \mathbf{c}_k \\ &= \sum_{k=1}^n \left(\frac{\text{SOMA}_{C=k}}{m} \right)^2 \mathbf{c}_k \cdot \mathbf{c}_k \\ &= \sum_{k=1}^n \frac{(\text{SOMA}_{C=k})^2}{m^2} m \\ &= \sum_{k=1}^n \frac{(\text{SOMA}_{C=k})^2}{m}. \end{aligned}$$

Assim, o quadrado médio para W_C é dado por:

$$QM(W_C) = \frac{\|P_{W_C}\mathbf{Y}\|^2}{\dim(W_C)} = \frac{\|P_{W_C}\mathbf{Y}\|^2}{n-1}.$$

A soma de quadrados para W_E , ou soma de quadrados dos resíduos, é obtida por diferença, isso porque como \mathbf{Y} é soma dos vetores $P_{V_0}\mathbf{Y}$, $P_{W_T}\mathbf{Y}$, $P_{W_L}\mathbf{Y}$, $P_{W_C}\mathbf{Y}$ e $P_{W_E}\mathbf{Y}$, temos que:

$$\|\mathbf{Y}\|^2 = \|P_{V_0}\mathbf{Y}\|^2 + \|P_{W_T}\mathbf{Y}\|^2 + \|P_{W_L}\mathbf{Y}\|^2 + \|P_{W_C}\mathbf{Y}\|^2 + \|P_{W_E}\mathbf{Y}\|^2.$$

Portanto a soma de quadrados para W_E é dada por:

$$\begin{aligned}
\|P_{W_E}\mathbf{Y}\|^2 &= \|\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 - \|P_{W_T}\mathbf{Y}\|^2 - \|P_{W_L}\mathbf{Y}\|^2 - \|P_{W_C}\mathbf{Y}\|^2 \\
&= \|\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 - (\|P_{V_T}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2) \\
&\quad - (\|P_{V_L}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2) - (\|P_{V_C}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2) \\
&= \|\mathbf{Y}\|^2 - \|P_{V_T}\mathbf{Y}\|^2 - \|P_{V_L}\mathbf{Y}\|^2 - \|P_{V_C}\mathbf{Y}\|^2 + 2(\|P_{V_0}\mathbf{Y}\|^2).
\end{aligned}$$

Definidas todas as fontes de variação, que são resultado da decomposição de V como soma direta, é possível se obter a análise de variância para o delineamento em linhas e colunas, assim como mostra a Tabela 3.9.

Tabela 3.9 – Tabela de análise de variância para delineamento em linhas e colunas.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	$QM(V_0)/QM(W_E)$
Tratamentos W_T	$t - 1$	$\ P_{W_T}\mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/QM(W_E)$
Linhas W_L	$m - 1$	$\ P_{W_L}\mathbf{Y}\ ^2$	$QM(W_L)$	$QM(W_L)/QM(W_E)$
Colunas W_C	$m - 1$	$\ P_{W_C}\mathbf{Y}\ ^2$	$QM(W_C)$	$QM(W_C)/QM(W_E)$
Resíduo W_E	$N - t - m - n + 2$	$\ P_{W_E}\mathbf{Y}\ ^2$	$QM(W_E)$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008)

Assuma que $\mathbf{Y} \sim N(\boldsymbol{\tau} + \boldsymbol{\zeta} + \boldsymbol{\eta}, \sigma^2\mathbf{I})$, onde $\boldsymbol{\tau} \in V_T$, $\boldsymbol{\zeta} \in V_L$ e $\boldsymbol{\eta} \in V_C$, de modo que o modelo aditivo para a soma de vetores é válido e que qualquer variação do experimento é decorrente do erro, em termos vetoriais:

$$\mathbf{Y} = \boldsymbol{\tau} + \boldsymbol{\zeta} + \boldsymbol{\eta} + \boldsymbol{\varepsilon},$$

em que $E(\mathbf{Y}) = \boldsymbol{\tau} + \boldsymbol{\zeta} + \boldsymbol{\eta} \in V_T + V_L + V_C$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ e $cov(\mathbf{Y}) = \sigma^2\mathbf{I}$.

Note que agora

$$\boldsymbol{\tau} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_T, \text{ com } \boldsymbol{\tau}_0 \in V_0 \text{ e } \boldsymbol{\tau}_T \in W_T, \text{ porque } V_T = V_0 + W_T;$$

$$\boldsymbol{\zeta} = \boldsymbol{\zeta}_0 + \boldsymbol{\zeta}_L, \text{ com } \boldsymbol{\zeta}_0 \in V_0 \text{ e } \boldsymbol{\zeta}_L \in W_L, \text{ porque } V_L = V_0 + W_L;$$

$$\boldsymbol{\eta} = \boldsymbol{\eta}_0 + \boldsymbol{\eta}_C, \text{ com } \boldsymbol{\eta}_0 \in V_0 \text{ e } \boldsymbol{\eta}_C \in W_C, \text{ porque } V_C = V_0 + W_C.$$

Portanto,

$$\begin{aligned} E(\mathbf{Y}) &= (\boldsymbol{\tau}_0 + \boldsymbol{\tau}_T) + (\boldsymbol{\zeta}_0 + \boldsymbol{\zeta}_L) + (\boldsymbol{\eta}_0 + \boldsymbol{\eta}_C) \\ &= (\boldsymbol{\tau}_0 + \boldsymbol{\zeta}_0 + \boldsymbol{\eta}_0) + (\boldsymbol{\tau}_T + \boldsymbol{\zeta}_L + \boldsymbol{\eta}_C). \end{aligned}$$

Como $\boldsymbol{\tau}_0$, $\boldsymbol{\zeta}_0$ e $\boldsymbol{\eta}_0$ estão todos em V_0 , são ambos múltiplo de \mathbf{u}_0 , sendo $\{\mathbf{u}_0\}$ base de V_0 . De modo que $\boldsymbol{\tau}_0$, $\boldsymbol{\zeta}_0$ e $\boldsymbol{\eta}_0$ não podem ser distinguidos pelos dados. Uma vez que os efeitos de tratamentos, linhas e colunas, encontram-se em W_T , W_L e W_C respectivamente, e esses são ortogonais entre si, é possível estimar combinações lineares dos efeitos, isso porque:

$$\boldsymbol{\tau}_T = \boldsymbol{\tau} - \boldsymbol{\tau}_0, \quad \boldsymbol{\zeta}_L = \boldsymbol{\zeta} - \boldsymbol{\zeta}_0 \text{ e } \boldsymbol{\eta}_C = \boldsymbol{\eta} - \boldsymbol{\eta}_0.$$

O teste de hipótese

Novamente, é possível fazer os testes de hipóteses, agora para quatro razões de variação. Os testes são feitos de forma análoga ao que feito nas Seções 3.1 e 3.2.

Lembre-se de que o teste para verificar se existe diferença entre as médias das linhas ou entre as médias das colunas nem sempre é de interesse do pesquisador. Já, o teste para verificar se existe diferença entre as médias dos tratamentos é o que permite inferir conclusões mais gerais sobre o experimento. As hipóteses a serem testadas para ambos os casos são respectivamente:

$$H_0 : \zeta_1 = \zeta_2 = \dots = \zeta_t$$

$$H_a : \boldsymbol{\zeta} \text{ não é um vetor constante.}$$

$$H_0 : \eta_1 = \eta_2 = \dots = \eta_t$$

$$H_a : \boldsymbol{\eta} \text{ não é um vetor constante.}$$

$$H_0 : \tau_1 = \tau_2 = \dots = \tau_t$$

$$H_a : \boldsymbol{\tau} \text{ não é um vetor constante.}$$

Para ambos os testes, proceda da mesma maneira ao que foi feito nas Seções 3.1 e 3.2. Compare cada vetor de projeção das fontes de variação com o vetor de resíduos. Para tanto, compare os quadrados médios pelas razões de variação e decida por não rejeitar ou rejeitar H_0 para cada teste acima.

Nesse momento, é necessário enfatizar algo que talvez o leitor não tenha notado. Os testes de hipóteses verificam a quais subespaços do modelo completo pertencem os vetores de médias. Sendo o modelo completo aquele que foi decomposto em soma direta e que não contém o subespaço residual. Desse modo, o teste é feito do menor subespaço para o maior. Como vimos, primeiramente testamos se o vetor de médias é um vetor constante, caso não seja, prosseguimos testando se o vetor de médias pertence a um subespaço maior.

Por exemplo, ao aceitar que $\boldsymbol{\tau}$ é um vetor constante, assumimos que $\boldsymbol{\tau} \in V_0$. Caso contrário, ao rejeitar que $\boldsymbol{\tau}$ é um vetor constante, assumimos que $\boldsymbol{\tau} \in V_T = V_0 \oplus W_T$.

No caso do delineamento em linhas e colunas, o modelo completo é $V_T + V_L + V_C$, as hipóteses são testadas para cada um dos subespaços dessa soma.

Exemplo prático

Para encerrar esta seção, é apresentado um exemplo de um experimento que deve ser conduzido no delineamento em linhas e colunas. Esse exemplo foi retirado de Bailey (2008).

Um experimento deseja comparar quatro tipos de vinhos. Os vinhos são provados e avaliados por oito juízes. Os juízes devem ser considerados como blocos, pois diferentes juízes podem dar notas discrepantes. Ao planejar o experimento, é necessário especificar a ordem de prova para cada juiz. Desse modo, a ordem das provas também deve ser considerada como blocos, isso porque, na quarta prova os juízes podem estar embriagados e, portanto dar boas notas aos vinhos ou, pelo contrário, estar fadigados, sensorialmente, e dar notas ruins.

Nesse exemplo, é necessário o uso de bloqueio duplo, onde a ordem das provas é considerada como linhas e o juizes como colunas. Temos, então, um experimento em linhas e colunas $4 \times 8 = 32$, com quatro tratamentos que são os vinhos A, B, C e D. A aleatorização dos vinhos às parcelas é feita como foi apresentado para o delineamento em linhas e colunas, para o caso geral. Veja a Tabela 3.10 e perceba que cada vinho é repetido duas vezes em cada linha e uma vez em cada coluna.

3.4 Experimento fatorial

No experimento fatorial os tratamentos consistem de todas as combinações dos níveis de dois ou mais fatores de tratamento². Será apresentado, nesta seção, o caso em que os tratamen-

² Os conceitos formais de níveis e fatores serão vistos no Capítulo 6.

Tabela 3.10 – Exemplo avaliação de vinhos.

Provas	J u í z e s							
	1	2	3	4	5	6	7	8
1	C	D	C	A	B	B	A	D
2	A	B	A	C	D	D	C	B
3	D	A	B	D	C	C	B	A
4	B	C	D	B	A	A	D	C

Fonte: Bailey (2008).

tos são combinações dos níveis de apenas dois fatores. Os procedimentos apresentados podem ser estendidos a três ou mais fatores de tratamento.

Segundo Pimentel-Gomes e Garcia (2002), experimentos fatoriais geralmente são mais eficientes do que experimentos sem combinações de níveis fatores e permitem tirar conclusões mais gerais. Entretanto, para os autores, a principal deficiência desse tipo de experimento é que o número de tratamentos aumenta rapidamente.

Experimentos fatoriais podem ser conduzidos nos três tipos de delineamentos apresentados neste trabalho, os quais são: inteiramente ao acaso, em blocos e em linhas e colunas. Por simplicidade, nesta seção o experimento fatorial é apresentado no delineamento inteiramente casualizado.

Suponha que os tratamentos sejam todas as combinações dos níveis de dois fatores F e G . Denote por n_F e n_G a quantidade de níveis de F e G respectivamente. Portanto, o número total de tratamentos no experimento fatorial é dado por $n_F \times n_G$. Veja a seguir os fatores F e G :

$$F = \{1, 2, \dots, n_F\}$$

$$G = \{1, 2, \dots, n_G\}.$$

Notação: Escreva $T = F \wedge G$ se os tratamentos são todas as combinações dos níveis dos fatores F e G .

Agora, a quantidade de parcelas do espaço de respostas V é dada por $n_F \times n_G \times r$, em que r é o número de repetições dos tratamentos. Novamente é possível definir os subespaços de V de forma análoga ao que foi feito nas Seções 3.1, 3.2 e 3.3.

O subespaço V_T

O subespaço V_T , que é o subespaço de tratamento, consiste naqueles vetores em V que são constantes em cada tratamento. Lembre-se de que agora os tratamentos são todas as combinações dos níveis de F e G . Então $\dim(V_T) = \dim(V_{F \wedge G}) = n_F \times n_G = t$.

Obtida a base $\{\mathbf{u}_1, \dots, \mathbf{u}_t\}$ para V_T , a projeção de \mathbf{Y} em V_T é dada por:

$$\begin{aligned}
 P_{V_T} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_t}{\mathbf{u}_t \cdot \mathbf{u}_t} \right) \mathbf{u}_t \\
 &= \sum_{i=1}^t \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\mathbf{u}_i \cdot \mathbf{u}_i} \right) \mathbf{u}_i \\
 &= \sum_{i=1}^t \left(\frac{\text{SOMA}_{T=i}}{r_i} \right) \mathbf{u}_i \\
 &= \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i \\
 &= \text{vetor de médias dos tratamentos.}
 \end{aligned}$$

Para se obter a análise de variância para o experimento fatorial, é preciso decompor o modelo completo onde as hipóteses serão testadas. Como já mencionado, o modelo completo são todas as decomposições em soma direta que não contêm o subespaço residual. Nesse caso, o delineamento é inteiramente casualizado, de modo que o modelo completo é V_T . Sendo assim, é necessário decompor esse subespaço. A decomposição de V_T leva à decomposição do vetor de média dos tratamentos, e conseqüentemente à decomposição da soma de quadrados. Entretanto, agora a decomposição de V_T tem certa peculiaridade e será mostrada a seguir.

Os subespaços V_F e V_G

O subespaço V_F , que é o subespaço dos níveis de F , consiste naqueles vetores em V que são constantes em cada nível de F . Então, $\dim(V_F) = n_F$.

O subespaço V_G , que é o subespaço dos níveis de G , consiste naqueles vetores em V que são constantes em cada nível de G . Então, $\dim(V_G) = n_G$.

Exemplo

Suponha um experimento fatorial 2×3 em que os tratamentos são todas as combinações dos níveis dos fatores $F = \{1, 2\}$ e $G = \{a, b, c\}$. Note que nesse exemplo $n_F = 2$, $n_G = 3$ e $\mathfrak{T} = \{1a, 1b, 1c, 2a, 2b, 2c\}$. Suponha que o número de repetições é igual para os seis tratamentos, $r = 2$. Logo, $N = 2 \times 3 \times 2 = 12$. Na Tabela 3.11, mostra-se o conjunto de parcelas Ω , o conjunto de tratamento \mathfrak{T} , os vetores $\mathbf{u}_{1a}, \mathbf{u}_{1b}, \mathbf{u}_{1c}, \mathbf{u}_{2a}, \mathbf{u}_{2b}$ e \mathbf{u}_{2c} da base para V_T , os vetores \mathbf{v}_1 e \mathbf{v}_2 da base para V_F e os vetores $\mathbf{v}_a, \mathbf{v}_b$ e \mathbf{v}_c da base para V_G .

Note que cada tratamento ocorre em r parcelas. Note também que cada nível de F ocorre em $n_G r$ parcelas, isso, porque $\mathbf{v}_j \cdot \mathbf{v}_j = n_G r$, sendo \mathbf{v}_j um vetor da base de V_F . Analogamente,

Tabela 3.11 – Vetores no experimento fatorial.

Ω	\mathfrak{T}	\mathbf{u}_{1a}	\mathbf{u}_{1b}	\mathbf{u}_{1c}	\mathbf{u}_{2a}	\mathbf{u}_{2b}	\mathbf{u}_{2c}	\mathbf{v}_1	\mathbf{v}_2	\mathbf{v}_a	\mathbf{v}_b	\mathbf{v}_c
1	1a	1	0	0	0	0	0	1	0	1	0	0
2	1c	0	0	1	0	0	0	1	0	0	0	1
3	2b	0	0	0	0	1	0	0	1	0	1	0
4	1a	1	0	0	0	0	0	1	0	1	0	0
5	1b	0	1	0	0	0	0	1	0	0	1	0
6	2b	0	0	0	0	1	0	0	1	0	1	0
7	2c	0	0	0	0	0	1	0	1	0	0	1
8	2c	0	0	0	0	0	1	0	1	0	0	1
9	2a	0	0	0	1	0	0	0	1	1	0	0
10	2a	0	0	0	1	0	0	0	1	1	0	0
11	1c	0	0	1	0	0	0	1	0	0	0	1
12	1b	0	1	0	0	0	0	1	0	0	1	0

Fonte: Da autora (2019).

cada nível de G ocorre em n_{Fr} parcelas, isso porque $\mathbf{v}_k \cdot \mathbf{v}_k = n_{Fr}$, sendo \mathbf{v}_k um vetor da base de V_G .

A partir das bases de V_F e V_G é possível obter as projeções de \mathbf{Y} nesses subespaços.

Seja $\{\mathbf{v}_1, \dots, \mathbf{v}_{n_F}\}$ uma base para V_F , então a projeção de \mathbf{Y} em V_F é dada por:

$$\begin{aligned}
 P_{V_F} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{v}_{n_F}}{\mathbf{v}_{n_F} \cdot \mathbf{v}_{n_F}} \right) \mathbf{v}_{n_F} \\
 &= \sum_{j=1}^{n_F} \left(\frac{\mathbf{Y} \cdot \mathbf{v}_j}{\mathbf{v}_j \cdot \mathbf{v}_j} \right) \mathbf{v}_j \\
 &= \sum_{j=1}^{n_F} \left(\frac{\text{SOMA}_{F=j}}{n_{Gr}} \right) \mathbf{v}_j \\
 &= \sum_{j=1}^{n_F} (\text{MÉDIA}_{F=j}) \mathbf{v}_j \\
 &= \text{vetor de médias dos níveis de F.}
 \end{aligned}$$

Seja $\{\mathbf{v}_1, \dots, \mathbf{v}_{n_G}\}$ uma base para V_G , então a projeção de \mathbf{Y} em V_G é dada por:

$$\begin{aligned}
 P_{V_G}\mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{v}_{n_G}}{\mathbf{v}_{n_G} \cdot \mathbf{v}_{n_G}} \right) \mathbf{v}_{n_G} \\
 &= \sum_{k=1}^{n_G} \left(\frac{\mathbf{Y} \cdot \mathbf{v}_k}{\mathbf{v}_k \cdot \mathbf{v}_k} \right) \mathbf{v}_k \\
 &= \sum_{k=1}^{n_G} \left(\frac{\text{SOMA}_{G=k}}{n_{FR}} \right) \mathbf{v}_k \\
 &= \sum_{k=1}^{n_G} (\text{MÉDIA}_{G=k}) \mathbf{v}_k \\
 &= \text{vetor de médias dos níveis de G.}
 \end{aligned}$$

Os subespaços W_F e W_G

O subespaço W_F , que é o subespaço de efeito dos níveis de F , é dado pelo conjunto de vetores:

$$W_F = V_F \cap V_0^\perp = \{\mathbf{v} \in V_F : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

De modo que subespaço W_F é constituído pelos vetores em V_F que são perpendiculares a V_0 . Assim V_F é decomposto em:

$$V_F = V_0 \oplus W_F.$$

A dimensão de W_F pode ser obtida por diferença:

$$\begin{aligned}
 \dim(W_F) &= \dim(V_F) - \dim(V_0) \\
 &= n_F - 1.
 \end{aligned}$$

O subespaço W_G , que é o subespaço de efeito dos níveis de G , é dado pelo conjunto de vetores:

$$W_G = V_G \cap V_0^\perp = \{\mathbf{v} \in V_G : \mathbf{v} \text{ é ortogonal a } V_0\}.$$

De modo que o subespaço W_G é constituído pelos vetores em V_G que são perpendiculares a V_0 . Assim V_G é decomposto em:

$$V_G = V_0 \oplus W_G.$$

A dimensão de W_G pode ser obtida por diferença:

$$\begin{aligned} \dim(W_G) &= \dim(V_G) - \dim(V_0) \\ &= n_G - 1. \end{aligned}$$

As decomposições de V_F e V_G em soma direta levam à decomposição dos vetores em \mathbf{Y} em W_F e W_G , respectivamente. Lembrando que $P_{V_0}\mathbf{Y} = \bar{Y}\mathbf{u}_0$, as projeções são:

$$\begin{aligned} P_{W_F}\mathbf{Y} &= P_{V_F}\mathbf{Y} - P_{V_0}\mathbf{Y} \\ &= \sum_{j=1}^{n_F} (\text{MÉDIA}_{F=j}) \mathbf{v}_j - \bar{Y}\mathbf{u}_0 \\ &= \text{vetor de médias dos níveis de F} - \text{vetor de média geral} \\ &= \text{vetor de efeitos dos níveis de F}, \end{aligned}$$

$$\begin{aligned} P_{W_G}\mathbf{Y} &= P_{V_G}\mathbf{Y} - P_{V_0}\mathbf{Y} \\ &= \sum_{k=1}^{n_G} (\text{MÉDIA}_{G=k}) \mathbf{v}_k - \bar{Y}\mathbf{u}_0 \\ &= \text{vetor de médias dos níveis de G} - \text{vetor de média geral} \\ &= \text{vetor de efeitos dos níveis de G}. \end{aligned}$$

O subespaço $W_{F \wedge G}$

O subespaço $W_{F \wedge G}$ é o subespaço de efeito da interação F por G . Com ele é possível verificar se existe diferença entre as médias das combinações dos níveis de F e G . Antes de se definir o subespaço $W_{F \wedge G}$ assumamos a condição do teorema que segue.

Teorema 3.4.1. *Se todas as combinações de níveis dos fatores (ou seja, todos os tratamentos) F e G ocorrem no mesmo número de parcelas, então os subespaços W_F e W_G são ortogonais.*

A demonstração desse teorema é análoga a do Teorema 3.2.1 e encontra-se em Bailey (2008). De posse do Teorema 3.4.1 é possível definir o subespaço:

$$\begin{aligned} V_F + V_G &= (V_0 \oplus W_F) + (V_0 \oplus W_G) \\ &= V_0 \oplus W_F \oplus W_G. \end{aligned}$$

Observe que $V_F + V_G \subset V_T$, então defina $W_{F \wedge G}$ da seguinte maneira:

$$W_{F \wedge G} = V_T \cap (V_F + V_G)^\perp = \{\mathbf{v} \in V_T : \mathbf{v} \text{ é ortogonal a } V_F + V_G\}.$$

De modo que $W_{F \wedge G}$ é constituído pelos vetores em V_T que são perpendiculares a $V_F + V_G$. Portanto, V_T é decomposto em:

$$\begin{aligned} V_T &= (V_F + V_G) \oplus W_{F \wedge G} \\ &= V_0 \oplus W_F \oplus W_G \oplus W_{F \wedge G}. \end{aligned}$$

A decomposição de V_T leva à decomposição de sua dimensão e também à decomposição do vetor $P_{V_T} \mathbf{Y}$, que é o vetor de média dos tratamentos. Na Tabela 3.12 mostram-se as decomposições de V_T para o modelo completo.

Tabela 3.12 – Decomposição de V_T para experimento fatorial.

Subespaços	V_T	=	V_0	⊕	W_F	⊕	W_G	⊕	$W_{F \wedge G}$
Dimensão	t	=	1	+	$(n_F - 1)$	+	$(n_G - 1)$	+	$(n_F - 1)(n_G - 1)$
Vetor	$P_{V_T} \mathbf{Y}$	=	$P_{V_0} \mathbf{Y}$	+	$P_{W_F} \mathbf{Y}$	+	$P_{W_G} \mathbf{Y}$	+	$P_{W_{F \wedge G}} \mathbf{Y}$

Fonte: Bailey (2008).

Então, a dimensão de $W_{F \wedge G}$ é obtida por diferença:

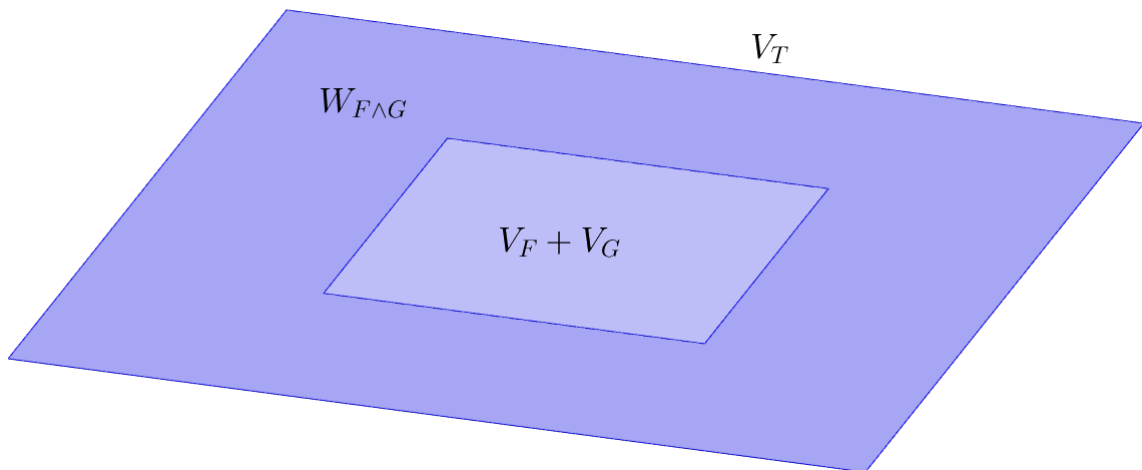
$$\begin{aligned} \dim(W_{F \wedge G}) &= \dim(V_T) - \dim(V_0) - \dim(W_F) - \dim(W_G) \\ &= t - 1 - (n_F - 1) - (n_G - 1) \\ &= (n_F - 1)(n_G - 1). \end{aligned}$$

A projeção de \mathbf{Y} em $W_{F \wedge G}$ também é obtida por diferença:

$$\begin{aligned}
 P_{W_{F \wedge G}} \mathbf{Y} &= P_{V_T} \mathbf{Y} - P_{V_0} \mathbf{Y} - P_{W_F} \mathbf{Y} - P_{W_G} \mathbf{Y} \\
 &= \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \bar{Y} \mathbf{u}_0 - \left(\sum_{j=1}^{n_F} (\text{MÉDIA}_{F=j}) \mathbf{v}_j - \bar{Y} \mathbf{u}_0 \right) \\
 &\quad - \left(\sum_{k=1}^{n_G} (\text{MÉDIA}_{G=k}) \mathbf{v}_k - \bar{Y} \mathbf{u}_0 \right) \\
 &= \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i - \sum_{j=1}^{n_F} (\text{MÉDIA}_{F=j}) \mathbf{v}_j - \sum_{k=1}^{n_G} (\text{MÉDIA}_{G=k}) \mathbf{v}_k + \bar{Y} \mathbf{u}_0 \\
 &= \text{vetor de médias dos tratamentos} - \text{vetor de médias dos níveis de F} \\
 &\quad - \text{vetor de médias dos níveis de G} + \text{vetor de média geral} \\
 &= \text{vetor de efeitos da interação F e G.}
 \end{aligned}$$

A Figura 3.15 pode ajudar a visualizar $V_F + V_G$ e $W_{F \wedge G}$ como subespaços de V_T .

Figura 3.15 – Subespaço V_T para o experimento fatorial.



Fonte: Da autora (2019).

O subespaço W_E

Os subespaços definidos para testar hipóteses são subespaços de V_T . Como aponta Bailey (2008), esses subespaços estão contidos em um modelo máximo V_T , que é utilizado para testar hipóteses sobre modelos e define o resíduo, de modo que, nada em V_T seja residual.

Dessa forma, é possível definir o subespaço W_E , também chamado de subespaço de resíduo para o experimento fatorial, que consiste no conjunto de vetores:

$$W_E = \{\mathbf{v} \in V : \mathbf{v} \text{ é ortogonal a } \mathbf{w} \text{ para todo } \mathbf{w} \in V_T\} = V_T^\perp.$$

Note que W_E é o complemento ortogonal de V_T . Desse modo, temos uma decomposição de espaço de resposta V em soma direta de:

$$V = V_T \oplus W_E.$$

A decomposição de V leva à decomposição de sua dimensão e também à decomposição do vetor \mathbf{Y} .

Então, é possível obter a dimensão de W_E por diferença, ou seja:

$$\begin{aligned} \dim(W_E) &= \dim(V) - \dim(V_T) \\ &= N - t. \end{aligned}$$

A projeção de \mathbf{Y} em W_E também é obtida por diferença, ou seja:

$$\begin{aligned} P_{W_E} \mathbf{Y} &= \mathbf{Y} - P_{V_T} \mathbf{Y} \\ &= \mathbf{Y} - \sum_{i=1}^t (\text{MÉDIA}_{T=i}) \mathbf{u}_i \\ &= \text{vetor de dados} - \text{vetor de médias dos tratamentos} . \end{aligned}$$

Para se obter as somas de quadrados e os quadrados médios para os subespaços das fontes de variação, use as quantidades associadas aos subespaços V , apresentados na Tabela 3.13.

Tabela 3.13 – Quantidades associadas aos subespaços V .

Subespaços	V_0	V_F	V_G	V_T
Dimensão	1	n_F	n_G	t
Vetor	$P_{V_0} \mathbf{Y}$	$P_{V_F} \mathbf{Y}$	$P_{V_G} \mathbf{Y}$	$P_{V_T} \mathbf{Y}$
$\ \text{Vetor}\ ^2$	$\frac{(\text{SOMA}_Y)^2}{N}$	$\sum_{j=1}^{n_F} \frac{(\text{SOMA}_{F=j})^2}{n_{Gr}}$	$\sum_{k=1}^{n_G} \frac{(\text{SOMA}_{G=k})^2}{n_{Fr}}$	$\sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i}$

Fonte: Adaptado Bailey (2008).

Os subespaços das fontes de variação são: V_0 , W_F , W_G , $W_{F \wedge G}$ e W_E . Para se obter as somas de quadrados e os quadrados médios associados a esses subespaços, basta subtrair as quantidades de somas de quadrados e os quadrados médios corretamente.

A soma de quadrados para V_0 , ou soma de quadrados para a média, é dada por:

$$\|P_{V_0}\mathbf{Y}\|^2 = \frac{(\text{SOMA}_Y)^2}{N}.$$

O quadrado médio para V_0 é dado por:

$$QM(V_0) = \frac{\|P_{V_0}\mathbf{Y}\|^2}{\dim(V_0)} = \frac{\|P_{V_0}\mathbf{Y}\|^2}{1}.$$

A soma de quadrados para W_F , ou soma de quadrados dos efeitos dos níveis de F , é dada por:

$$\begin{aligned} \|P_{W_F}\mathbf{Y}\|^2 &= \|P_{V_F}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 \\ &= \sum_{j=1}^{n_F} \frac{(\text{SOMA}_{F=j})^2}{n_{Gr}} - \frac{(\text{SOMA}_Y)^2}{N}. \end{aligned}$$

O quadrado médio para W_F é dado por:

$$QM(W_F) = \frac{\|P_{W_F}\mathbf{Y}\|^2}{\dim(W_F)} = \frac{\|P_{W_F}\mathbf{Y}\|^2}{n_F - 1}.$$

A soma de quadrados para W_G , ou soma de quadrados dos efeitos dos níveis de G , é dada por:

$$\begin{aligned} \|P_{W_G}\mathbf{Y}\|^2 &= \|P_{V_G}\mathbf{Y}\|^2 - \|P_{V_0}\mathbf{Y}\|^2 \\ &= \sum_{k=1}^{n_G} \frac{(\text{SOMA}_{G=k})^2}{n_{Fr}} - \frac{(\text{SOMA}_Y)^2}{N}. \end{aligned}$$

O quadrado médio para W_G é dado por:

$$QM(W_G) = \frac{\|P_{W_G}\mathbf{Y}\|^2}{\dim(W_G)} = \frac{\|P_{W_G}\mathbf{Y}\|^2}{n_G - 1}.$$

Nota-se que a soma de quadrados para V_T , ou soma de quadrados das médias dos tratamentos, é decomposta em média geral, efeitos dos níveis de F , efeitos dos níveis de G e efeitos

da interação F e G . A soma de quadrados para $W_{F \wedge G}$, ou soma de quadrados dos efeitos da interação, pode ser obtida por diferença como segue:

$$\begin{aligned}
\|P_{W_{F \wedge G}} \mathbf{Y}\|^2 &= \|P_{V_T} \mathbf{Y}\|^2 - \|P_{V_0} \mathbf{Y}\|^2 - \|P_{W_F} \mathbf{Y}\|^2 - \|P_{W_G} \mathbf{Y}\|^2 \\
&= \sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i} - \frac{(\text{SOMA}_Y)^2}{N} - \left(\sum_{j=1}^{n_F} \frac{(\text{SOMA}_{F=j})^2}{n_{Gr}} - \frac{(\text{SOMA}_Y)^2}{N} \right) \\
&\quad - \left(\sum_{k=1}^{n_G} \frac{(\text{SOMA}_{G=k})^2}{n_{Fr}} - \frac{(\text{SOMA}_Y)^2}{N} \right) \\
&= \sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i} - \sum_{j=1}^{n_F} \frac{(\text{SOMA}_{F=j})^2}{n_{Gr}} \\
&\quad - \sum_{k=1}^{n_G} \frac{(\text{SOMA}_{G=k})^2}{n_{Fr}} + \frac{(\text{SOMA}_Y)^2}{N}.
\end{aligned}$$

O quadrado médio para $W_{F \wedge G}$ é dado por:

$$QM(W_{F \wedge G}) = \frac{\|P_{W_{F \wedge G}} \mathbf{Y}\|^2}{\dim(W_{F \wedge G})} = \frac{\|P_{W_{F \wedge G}} \mathbf{Y}\|^2}{(n_F - 1)(n_G - 1)}.$$

Para a última fonte de variação, que é o resíduo, como $V = V_T \oplus W_E$, tem-se que a soma de quadrados para W_E é obtida por diferença:

$$\begin{aligned}
\|P_{W_E} \mathbf{Y}\|^2 &= \|\mathbf{Y}\|^2 - \|P_{V_T} \mathbf{Y}\|^2 \\
&= \|\mathbf{Y}\|^2 - \sum_{i=1}^t \frac{(\text{SOMA}_{T=i})^2}{r_i},
\end{aligned}$$

sendo o quadrado médio para W_E é dado por:

$$QM(W_E) = \frac{\|P_{W_E} \mathbf{Y}\|^2}{\dim(W_E)} = \frac{\|P_{W_E} \mathbf{Y}\|^2}{N - t}.$$

Agora, é possível se obter a análise de variância para o experimento fatorial como se mostra na Tabela 3.14. De acordo com Bailey (2008), quando os tratamentos são todas as combinações dos níveis dos fatores F e G , a linha para os tratamentos é substituída por efeitos dos níveis de F e G e efeitos da interação, de acordo com a decomposição de V_T .

Novamente, assumo que $\mathbf{Y} \sim N(\boldsymbol{\tau}, \sigma^2 \mathbf{I})$ onde $\boldsymbol{\tau} \in V_T$, de modo que o modelo aditivo para soma de vetores é correto e que qualquer variação no experimento é devida ao erro. Em

Tabela 3.14 – Tabela de análise de variância para experimento fatorial.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	$QM(V_0)/QM(W_E)$
Níveis de F W_F	$n_F - 1$	$\ P_{W_F}\mathbf{Y}\ ^2$	$QM(W_F)$	$QM(W_F)/QM(W_E)$
Níveis de G W_G	$n_G - 1$	$\ P_{W_G}\mathbf{Y}\ ^2$	$QM(W_G)$	$QM(W_G)/QM(W_E)$
Interação F por G $W_{F \wedge G}$	$(n_F - 1)(n_G - 1)$	$\ P_{W_{F \wedge G}}\mathbf{Y}\ ^2$	$QM(W_{F \wedge G})$	$QM(W_{F \wedge G})/QM(W_E)$
Resíduo W_E	$N - t$	$\ P_{W_E}\mathbf{Y}\ ^2$	$QM(W_E)$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

termos vetoriais temos:

$$\mathbf{Y} = \boldsymbol{\tau} + \boldsymbol{\varepsilon},$$

em que $E(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$, $E(\boldsymbol{\varepsilon}) = \mathbf{0}$ e $cov(\mathbf{Y}) = \sigma^2 \mathbf{I}$.

Para o experimento fatorial, os possíveis modelos para valor esperado do vetor de respostas são:

- $E(\mathbf{Y}) = \boldsymbol{\tau}_T \in V_T$, modelo de tratamento completo;
- $E(\mathbf{Y}) = \boldsymbol{\tau}_{F+G} \in V_F + V_G$, modelo aditivo em F e G ;
- $E(\mathbf{Y}) = \boldsymbol{\tau}_F \in V_F$, modelo em F apenas;
- $E(\mathbf{Y}) = \boldsymbol{\tau}_G \in V_G$, modelo em G apenas;
- $E(\mathbf{Y}) = \boldsymbol{\tau}_0 \in V_0$, modelo constante.

Se $E(\mathbf{Y}) \in V_F + V_G$, a interação F por G é nula, porque o subespaço da interação é $W_{F \wedge G} = V_T \cap (V_F + V_G)^\perp$. O que significa que os níveis de F e G são independentes. Sendo assim, é possível testar os efeitos dos tratamentos de F e G separadamente. Para o caso em que W_F e W_G são ortogonais e a interação F por G é nula, a equação $V_F + V_G = V_0 \oplus W_F \oplus W_G$ mostra que:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_{F+G} = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_F + \boldsymbol{\tau}_G.$$

Se $E(\mathbf{Y}) \in V_T$, a interação F por G é não nula, porque $V_T = (V_F + V_G) \oplus W_{F \wedge G}$. O que significa que os níveis de F e G são não independentes. Sendo assim, é necessário testar os

efeitos das combinações entre os tratamentos de F e G . Para caso que W_F e W_G são ortogonais e a interação F por G é não nula, a equação $V_T = V_0 \oplus W_F \oplus W_G \oplus W_{F \wedge G}$ mostra que:

$$\boldsymbol{\tau} = \boldsymbol{\tau}_T = \boldsymbol{\tau}_0 + \boldsymbol{\tau}_F + \boldsymbol{\tau}_G + \boldsymbol{\tau}_{F \wedge G}$$

e, então, a interação F por G é dada por:

$$\boldsymbol{\tau}_{F \wedge G} = \boldsymbol{\tau}_T - \boldsymbol{\tau}_0 - \boldsymbol{\tau}_F - \boldsymbol{\tau}_G.$$

De acordo com Bailey (2008), $\boldsymbol{\tau}_{F \wedge G}$ é uma medida de saída do ajuste do submodelo $V_F + V_G$ para o modelo completo V_T .

O teste de hipótese

Para realizar o teste de hipótese, é necessário verificar a qual subespaço pertence o vetor de médias do experimento. Para tanto, obtenha as hipóteses e compare os quadrados médios pelas razões de variação.

Como os testes são feitos do menor subespaço para o maior. Primeiramente, teste se existe efeito de interação F por G pelas hipóteses:

$$H_0 : \text{interação nula}$$

$$H_a : \text{interação não nula.}$$

Se a hipótese nula é verdadeira, então $\boldsymbol{\tau} \in V_F + V_G$ e, portanto, $\boldsymbol{\tau} = \boldsymbol{\tau}_{F+G}$. Se a hipótese nula é falsa, então $\boldsymbol{\tau} \in V_T$ e portanto $\boldsymbol{\tau} = \boldsymbol{\tau}_T$. Sob H_0 verdadeira a nuvem de pontos é centrada em $\boldsymbol{\tau}_{F+G}$ que pertence ao subespaço $V_F + V_G$, como se mostra na Figura 3.16.

Para realizar o teste, compare o vetor $P_{W_{F \wedge G}} \mathbf{Y}$ com o vetor de resíduos $P_{W_E} \mathbf{Y}$ (Figura 3.16). Para tanto, use a razão de variâncias $QM(W_{F \wedge G})/QM(W_E)$. Se esse valor for menor que um, não se rejeita H_0 e pode-se concluir que a interação é nula, logo, os efeitos dos níveis de F e G são independentes. Se esse valor for maior que um, rejeita-se H_0 e pode-se concluir que a interação é não nula, logo, os efeitos dos níveis de F e G são dependentes.

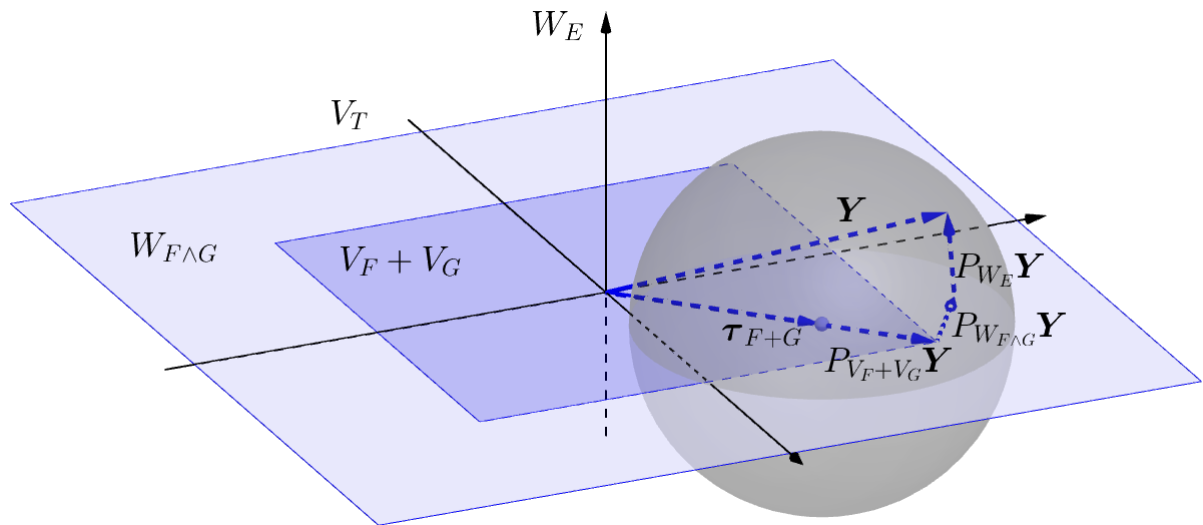
Para o caso em que a interação é nula, é necessário testar os efeitos dos níveis de F e G , separadamente, pelas hipóteses que seguem.

Para testar os efeitos dos níveis de F use as hipóteses:

$$H_0 : \boldsymbol{\tau}_{F_1} = \boldsymbol{\tau}_{F_2} = \dots = \boldsymbol{\tau}_{F_{n_F}}$$

$$H_a : \boldsymbol{\tau}_F \text{ não é um vetor constante.}$$

Figura 3.16 – Teste de hipótese para o experimento fatorial.



Fonte: Da autora (2019).

Se $QM(W_F)/QM(W_E)$ for menor que um, não rejeite H_0 e conclua que τ_F é um vetor constante, logo, não existe diferença entre as médias dos níveis de F .

Para testar os efeitos dos níveis de G use as hipóteses:

$$H_0 : \tau_{G_1} = \tau_{G_2} = \dots = \tau_{G_{n_G}}$$

$$H_a : \tau_G \text{ não é um vetor constante.}$$

Se $QM(W_G)/QM(W_E)$ for menor que um, não rejeite H_0 e conclua que τ_G é um vetor constante, logo, não existe diferença entre as médias dos níveis de G .

Na análise usual, o efeito de interação é estudado, plotando-se as médias das respostas do experimento em um gráfico. O exemplo, a seguir, mostra como proceder.

Exemplo para análise no experimento fatorial

Para realizar a análise, é possível plotar as médias das respostas para cada nível do fator F e manter fixos os níveis do fator G . Nesse caso, obtêm-se curvas para diferentes níveis de G . Se as curvas são paralelas, então, a interação é nula. Se as curvas não são paralelas, ou até mesmo, se cruzam, a interação é não nula (ADÃO, 2011).

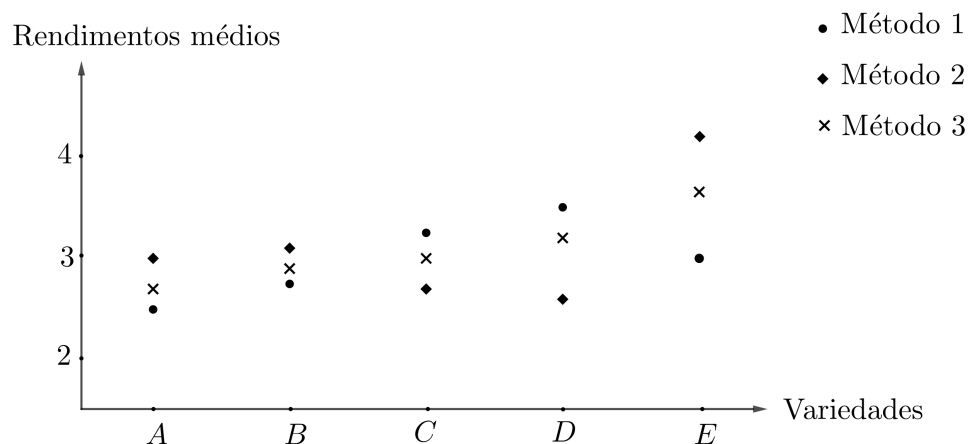
Da mesma forma, é possível plotar as médias das respostas para cada nível do fator G e manter fixo os níveis do fator F . Nesse caso, obtêm-se curvas para diferentes níveis de F .

Caso haja um fator de tratamento com níveis qualitativos e outro com níveis quantitativos, plote as médias dos níveis qualitativos e mantenha fixos os níveis quantitativos.

O exemplo, a seguir, retirado de Bailey (2008), pode esclarecer como a análise é realizada.

Em um experimento os tratamentos consistem de cinco variedades de ervilha combinadas com três métodos de cultivo. Desse modo, o conjunto das variedades é $F = \{A, B, C, D, E\}$, e o conjunto dos métodos é $G = \{1, 2, 3\}$. Para realizar a análise, os rendimentos médios, em toneladas por hectare, são mostrados na Figura 3.17. Nesse caso, foram plotadas as médias para os métodos mantendo fixas as variedades.

Figura 3.17 – Interação entre variedades e métodos de cultivo.



Fonte: Bailey (2008).

Na Figura 3.17, mostra-se que existe interação, uma vez que, as curvas para diferentes variedades de ervilha não são paralelas e se cruzam. Note que o método de cultivo 1 é melhor para as variedades C e D e pior para A, B e E. Já o método de cultivo 2 é melhor para as variedades A, B e E e pior para C e D. Enquanto o método de cultivo 3 é intermediário para todas as variedades.

Conduzui-se, este experimento, com o objetivo de mostrar qual rendimento médio de produção é melhor em cada método, considerando todas as variedades. O agricultor com sua experiência, de posse desses resultados, provavelmente deve saber qual combinação é a mais vantajosa para ser utilizada em campo.

3.5 Princípios de modelos de esperança

O subespaço onde se assume pertencer o valor esperado de um vetor de variáveis aleatórias é chamado de modelo de esperança. Como visto em capítulos anteriores, o valor esperado do

vetor aleatório \mathbf{Y} pode pertencer a vários subespaços. Entretanto, até o presente momento, não se definiu como uma coleção de subespaços de V pode ser utilizada para testar hipóteses.

Como aponta Bailey (2008), o teste de hipótese é feito para selecionar qual o menor modelo suportado pelos dados e, em seguida, estimar os parâmetros desse modelo. De modo que uma coleção de subespaços de V não pode ser arbitrária e deve satisfazer a alguns princípios. Esses princípios são apresentados a seguir, conforme constam em Bailey (2008).

Princípio 3.5.1. Princípio da Intersecção: *Se V_1 e V_2 são ambos modelos de esperança então $V_1 \cap V_2$ também deve ser um modelo de esperança.*

Nota-se que, ao definir $V_1 \cap V_2$ como modelo de esperança, não é necessário escolher entre V_1 ou V_2 , caso o valor estimado pertença a $V_1 \cap V_2$. De modo que, se \mathbf{Y} está em $V_1 \cap V_2$ então o modelo ajustado deve ser V_1 e também V_2 .

Princípio 3.5.2. Princípio da Soma: *Se V_1 e V_2 são ambos modelos de esperança então $V_1 + V_2$ também deve ser um modelo de esperança.*

A soma de subespaços é uma característica de todos os modelos lineares para permitir a soma de vetores. Novamente, ao definir $V_1 + V_2$ como modelo de esperança, não é necessário escolher entre V_1 ou V_2 , caso o valor estimado pertença a $V_1 + V_2$.

Todos os modelos de esperança estão contidos em um único modelo máximo. Esse modelo é o ponto de partida para testar hipóteses sobre modelos e define o resíduo. De modo que, nada no modelo máximo é residual, mesmo quando encaixado em um modelo menor. Nas Seções 3.1 e 3.4 o modelo máximo é V_T , na Seção 3.2 o modelo máximo é $V_T + V_B$ e na Seção 3.3 o modelo máximo é $V_T + V_L + V_C$. Portanto, é fundamental que a soma de subespaços seja um modelo de esperança.

Princípio 3.5.3. Princípio da Ortogonalidade: *Se V_1 e V_2 são ambos modelos de esperança e se $W = V_1 \cap V_2$ então $V_1 \cap W^\perp$ deve ser ortogonal a $V_2 \cap W^\perp$.*

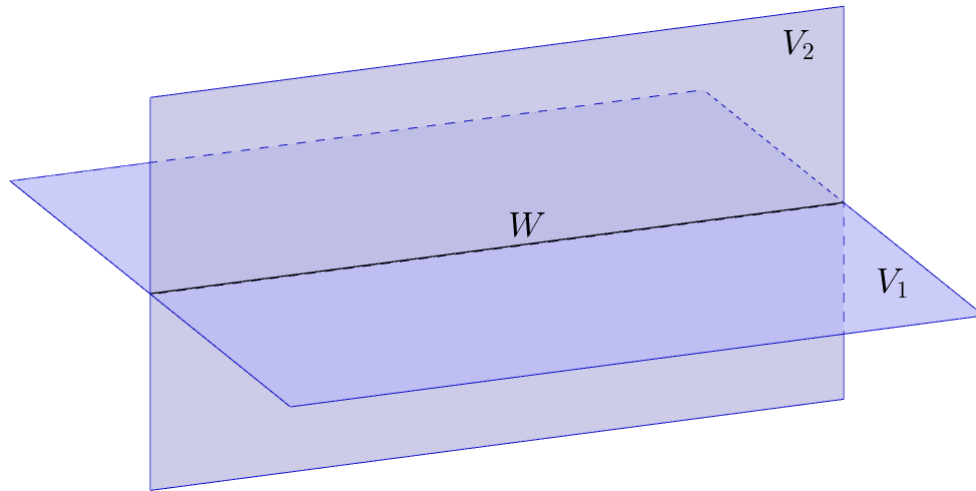
Dados os modelos de esperança V_1 e V_2 e os subespaços $W = V_1 \cap V_2$, $V_1 \cap W^\perp$ e $V_2 \cap W^\perp$, temos que:

$$V_1 = W + V_1 \cap W^\perp \text{ e}$$

$$V_2 = W + V_2 \cap W^\perp.$$

Logo, $V_1 + V_2 = W + V_1 \cap W^\perp + V_2 \cap W^\perp$.

Figura 3.18 – Subespaços para o princípio da ortogonalidade.



Fonte: Da autora (2019).

Suponha que a hipótese de que $E(\mathbf{Y}) \in V_1 + V_2$ é verdadeira. Desse modo, pode-se verificar, se é possível reduzir o modelo de esperança para V_1 . Para tanto, examine o tamanho do vetor $P_{V_1+V_2}\mathbf{Y} - P_{V_1}\mathbf{Y}$.

Como

$$\begin{aligned} \dim(V_1 + V_2) - \dim(V_1) &= (\dim(V_1) + \dim(V_2) - \dim(V_1 \cap V_2)) - \dim(V_1) \\ &= \dim(V_2) - \dim(W) \\ &= \dim(V_2 \cap W^\perp), \end{aligned}$$

temos que $P_{V_1+V_2}\mathbf{Y} - P_{V_1}\mathbf{Y}$ é igual a $P_{V_2 \cap W^\perp}\mathbf{Y}$, se $V_1 \cap W^\perp$ é ortogonal a $V_2 \cap W^\perp$. Para examinar o tamanho do vetor $P_{V_1+V_2}\mathbf{Y} - P_{V_1}\mathbf{Y}$, compare o quadrado médio

$$\frac{\|P_{V_1+V_2}\mathbf{Y} - P_{V_1}\mathbf{Y}\|^2}{\dim(V_1 + V_2) - \dim(V_1)} = \frac{\|P_{V_2 \cap W^\perp}\mathbf{Y}\|^2}{\dim(V_2) - \dim(W)},$$

com o quadrado médio residual. Se esse valor for menor que um, a hipótese que $E(\mathbf{Y}) \in V_1$ é verdadeira. Então, pode-se prosseguir e testar se é possível reduzir o modelo para $W = V_1 \cap V_2$. Da mesma forma, examine o tamanho do vetor $P_{V_1}\mathbf{Y} - P_W\mathbf{Y}$.

Como

$$\dim(V_1) - \dim(W) = \dim(V_1 \cap W^\perp),$$

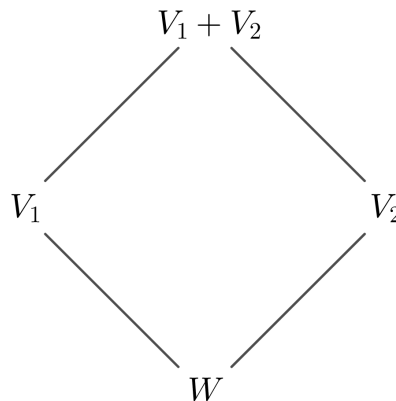
temos que $P_{V_1}\mathbf{Y} - P_W\mathbf{Y}$ é igual a $P_{V_1 \cap W^\perp}\mathbf{Y}$, se $V_1 \cap W^\perp$ é ortogonal a $V_2 \cap W^\perp$. Para examinar o tamanho do vetor $P_{V_1}\mathbf{Y} - P_W\mathbf{Y}$, compare o quadrado médio

$$\frac{\|P_{V_1}\mathbf{Y} - P_W\mathbf{Y}\|^2}{\dim(V_1) - \dim(W)} = \frac{\|P_{V_1 \cap W^\perp}\mathbf{Y}\|^2}{\dim(V_1) - \dim(W)},$$

com o quadrado médio residual. Se esse valor for menor que um, a hipótese que $E(\mathbf{Y}) \in W$ é verdadeira. Como W é o menor modelo suportado pelos dados os testes podem ser encerrados.

Note que o teste para reduzir o modelo de $V_1 + V_2$ a V_1 é o mesmo para reduzir o modelo de $V_1 + V_2$ a V_2 e, posteriormente, a W , de modo que, existem dois caminhos para W conforme pode ser visto na Figura 3.19. Pela ortogonalidade de $V_1 \cap W^\perp$ e $V_2 \cap W^\perp$, ambos os caminhos estão corretos porque em ambos os casos foi examinado $P_{V_1 \cap W^\perp}\mathbf{Y}$ e $P_{V_2 \cap W^\perp}\mathbf{Y}$. Se $V_1 \cap W^\perp$ e $V_2 \cap W^\perp$ não são ortogonais, então alguns valores de \mathbf{Y} fornecem resultados contraditórios e a inferência fica comprometida (BAILEY, 2008).

Figura 3.19 – Submodelos encaixados de $V_1 + V_2$.



Fonte: Adaptado Bailey (2008).

O princípio da ortogonalidade se aplica ao delineamento em blocos casualizados, delineamentos em linhas e colunas e experimentos fatoriais. Foi nesses casos, que se definiu a soma de dois subespaços como modelo de esperança. Para que esses delineamentos satisfaçam o princípio da ortogonalidade, é necessário garantir a condição do Teorema 3.4.1. Desse modo, o princípio da ortogonalidade é afetado pelo número de repetições dos tratamentos. Sendo assim, o princípio da ortogonalidade tem implicações no planejamento do experimento.

Em resumo,

Dada uma coleção de modelos de esperança que satisfaz os três princípios, testamos os submodelos, iniciando no modelo máximo e trabalhando para modelos menores. Em cada estágio, testamos o próximo submodelo examinando

a diferença entre as somas de quadrados para o ajuste no modelo atual e para o ajuste do submodelos dividido pela diferença entre suas dimensões. Esse quadrado médio é sempre comparado com o quadrado médio residual original. Se aceitamos o submodelo, descemos o diagrama até ele e continuamos a partir dele. Se em algum momento rejeitamos todos os submodelos imediatamente abaixo do modelo atual, decidimos que o atual é o menor suportado pelos dados. Devido à ortogonalidade, não importa em que ordem testamos os submodelos quando há uma escolha (BAILEY, 2008, p. 86, tradução nossa).

4 RELAÇÕES ENTRE A GEOMETRIA DOS MODELOS LINEARES E A GEOMETRIA DA ANÁLISE DE VARIÂNCIA

Como aponta Bailey (2008), alguns autores usam a notação matricial para expressar modelos lineares. Essa notação admite uma interpretação geométrica.

Neste capítulo objetivou-se fazer uma comparação entre a abordagem utilizada no presente trabalho com a abordagem usual empregada em modelos lineares.

4.1 Modelos lineares na presença de efeitos de tratamentos

O modelo para o delineamento inteiramente casualizado, na notação matricial, considera a matriz de delineamento $\mathbf{X}_{N \times t}$ e o vetor de médias dos tratamentos $\boldsymbol{\beta}_{t \times 1}$ definidos por:

$$\mathbf{X} = \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix} \text{ e } \boldsymbol{\beta} = \begin{pmatrix} \tau_1 \\ \tau_2 \\ \cdots \\ \tau_t \end{pmatrix},$$

sendo $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_t$ vetores coluna e $\tau_1, \tau_2, \dots, \tau_t$ as médias dos tratamentos. Note que, nesse caso, a matriz \mathbf{X} não possui a coluna correspondente ao vetor \mathbf{u}_0 .

Matricialmente, o modelo linear é expresso como:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

em que \mathbf{Y} é o vetor de dados, \mathbf{X} a matriz de delineamento, $\boldsymbol{\beta}$ é o vetor de efeitos de tratamentos e $\boldsymbol{\varepsilon}$ é o vetor de resíduos.

Assim como em Bailey (2008), vale destacar que $\boldsymbol{\beta}$ não é o mesmo que $\boldsymbol{\tau}$, definido na Seção 3.1, isso porque $\boldsymbol{\beta}_{t \times 1}$, enquanto $\boldsymbol{\tau}_{N \times 1}$. De modo que $\boldsymbol{\tau} = \mathbf{X}\boldsymbol{\beta}$, ou seja, $\boldsymbol{\tau}$ contém coordenadas relativas às repetições dos tratamentos.

Na abordagem geométrica, a teoria dos modelos lineares, se caracteriza por considerar a matriz de delineamento \mathbf{X} como uma transformação linear ¹ do espaço de parâmetros \mathbb{R}^t no espaço de dados \mathbb{R}^N . A transformação linear \mathbf{X} aplicada a um vetor no espaço de parâmetros resulta em um vetor em imagem \mathbf{X} . Dessa forma, a imagem de \mathbf{X} é o modelo linear para as

¹ Veja Definição .0.8 do Apêndice deste trabalho.

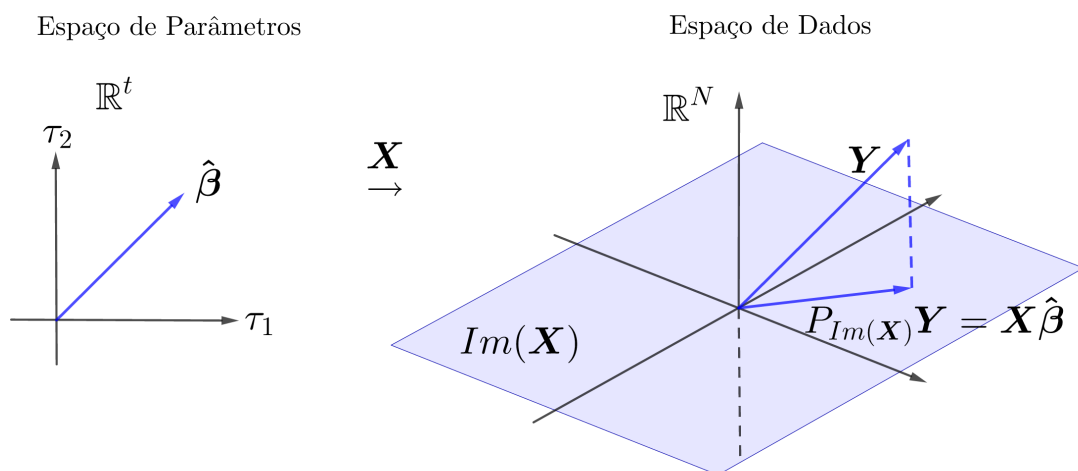
médias do vetor aleatório \mathbf{Y} . Assim, o subespaço imagem de \mathbf{X} contém todos os vetores que podem representar o valor esperado de \mathbf{Y} , ou seja, o vetor $\boldsymbol{\tau} = \mathbf{X}\boldsymbol{\beta}$.

Uma das soluções para o modelo matricial é obtida pelo método dos mínimos quadrados, que consiste, simplesmente, em obter um vetor em imagem de \mathbf{X} o mais próximo possível do vetor de dados \mathbf{Y} . Isso é obtido pela projeção ortogonal de \mathbf{Y} em imagem de \mathbf{X} , ou seja, $P_{Im(\mathbf{X})}\mathbf{Y}$. Desse que, o processo de estimação pode ser descrito como uma regra que atribui a cada vetor observado \mathbf{Y} um vetor em $Im(\mathbf{X})$. A estimativa de mínimos quadrados é, então, obtida pela pré-imagem de \mathbf{X} , considerada aqui de posto máximo e, portanto, injetiva ($N > p$), isto é:

$$\mathbf{X}\hat{\boldsymbol{\beta}} = P_{Im(\mathbf{X})}\mathbf{Y},$$

e como \mathbf{X} é injetiva, existe apenas um vetor $\hat{\boldsymbol{\beta}}$ em \mathbb{R}^t que satisfaz a equação acima. Note que $P_{Im(\mathbf{X})}\mathbf{Y}$ é o vetor de médias ajustadas dos tratamentos ou vetor de dados ajustados $\hat{\mathbf{Y}}$.

Figura 4.1 – Matriz \mathbf{X} como transformação linear.



Fonte: Da autora (2019).

Em termos matriciais, a projeção ortogonal é dada por:

$$P_{Im(\mathbf{X})}\mathbf{Y} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}$$

$$\mathbf{X}\hat{\boldsymbol{\beta}} = \mathbf{X}(\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y},$$

e pela injetividade da \mathbf{X} temos que:

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}^T\mathbf{X})^{-1}\mathbf{X}^T\mathbf{Y}.$$

Como \mathbf{X} não possui a coluna correspondente a \mathbf{u}_0 , $(\mathbf{X}^\top \mathbf{X})^{-1}$ possui inversa simples.

Na análise de variância, para o delineamento inteiramente casualizado, assumimos que os valores observados assumem o modelo

$$Y_{ij} = \tau_i + \varepsilon_{ij},$$

para $i = 1, \dots, t$ (t número de tratamentos) e $j = 1, \dots, r_i$ (r_i número de repetições do tratamento i).

As coordenadas do espaço de dados \mathbb{R}^N são parametrizadas pelas parcelas. Essa parametrização é considerada como a base para o subespaço de tratamentos V_T . As coordenadas de um vetor \mathbf{u}_i da base são:

$$\begin{cases} 1, & \text{se o tratamento } i \text{ foi aplicado à parcela } \omega \\ 0, & \text{caso contrário.} \end{cases}$$

De modo que a base para V_T é dada por $\{\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_t\}$.

Definindo o espaço de parâmetros como $\mathbb{R}^t = \{(\tau_1, \tau_2, \dots, \tau_t), \tau_i \in \mathbb{R}\}$ e a base canônica para esse espaço $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_t\}$. Dessa forma, a matriz \mathbf{X} pode ser definida por $\mathbf{X}\mathbf{e}_i = \mathbf{u}_i$, de onde vem que a matriz \mathbf{X} é dada por:

$$\begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix}.$$

Dessa forma, a imagem de \mathbf{X} é exatamente o subespaço de tratamentos V_T .

Observe que \mathbf{X} tem posto completo e

$$\mathbf{X}^\top \mathbf{X} = \begin{pmatrix} r_1 & 0 & \cdots & 0 \\ 0 & r_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & r_t \end{pmatrix}$$

tem inversa. Então, a projeção é simplesmente:

$$\begin{aligned} \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top &= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix} \begin{pmatrix} 1/r_1 & 0 & \cdots & 0 \\ 0 & 1/r_2 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & \cdots & 1/r_t \end{pmatrix} \begin{pmatrix} \mathbf{u}_1^\top \\ \mathbf{u}_2^\top \\ \cdots \\ \mathbf{u}_t^\top \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix} \begin{pmatrix} \frac{1}{r_1} \mathbf{u}_1^\top \\ \frac{1}{r_2} \mathbf{u}_2^\top \\ \cdots \\ \frac{1}{r_t} \mathbf{u}_t^\top \end{pmatrix}. \end{aligned}$$

Portanto,

$$\begin{aligned} \mathbf{X}(\mathbf{X}^\top \mathbf{X})^{-1} \mathbf{X}^\top \mathbf{Y} &= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix} \begin{pmatrix} \frac{1}{r_1} \mathbf{u}_1^\top \\ \frac{1}{r_2} \mathbf{u}_2^\top \\ \cdots \\ \frac{1}{r_t} \mathbf{u}_t^\top \end{pmatrix} \begin{pmatrix} Y_1 \\ Y_2 \\ \cdots \\ Y_N \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_t \end{pmatrix} \begin{pmatrix} \frac{1}{r_1} \mathbf{u}_1^\top \mathbf{Y} \\ \frac{1}{r_2} \mathbf{u}_2^\top \mathbf{Y} \\ \cdots \\ \frac{1}{r_t} \mathbf{u}_t^\top \mathbf{Y} \end{pmatrix} \\ &= \begin{pmatrix} \frac{1}{r_1} \mathbf{u}_1^\top \mathbf{Y} \\ \frac{1}{r_2} \mathbf{u}_2^\top \mathbf{Y} \\ \cdots \\ \frac{1}{r_t} \mathbf{u}_t^\top \mathbf{Y} \end{pmatrix}. \end{aligned}$$

Como $\frac{1}{r_i} \mathbf{u}_i^\top \mathbf{Y}$ é a estimativa do parâmetro de tratamento τ_i . Desse modo, o parâmetro de tratamento τ_i encontra-se nas coordenadas do vetor $\tau_i \mathbf{u}_i$, de modo que não há necessidade de explicitar o espaço paramétrico \mathbb{R}^t e toda análise de variância é realizada no espaço de dados \mathbb{R}^N .

Outro tipo de parametrização

É comum no delineamento inteiramente casualizado, na notação matricial, se trabalhar com a matriz de delineamento, da seguinte maneira

$$\mathbf{X} = \begin{pmatrix} \mathbf{u}_0 & \mathbf{u}_1 & \cdots & \mathbf{u}_t \end{pmatrix}.$$

Nesse caso, temos outro tipo de parametrização, de modo que o espaço de parâmetros é \mathbb{R}^{t+1} e o espaço de dados é \mathbb{R}^N . Agora \mathbf{X} não tem mais posto completo e o processo de estimação demanda inversas generalizadas, isso, porque $\mathbf{X}^T \mathbf{X}$ não tem inversa simples.

Na análise de variância, para essa parametrização, é comum assumir que os valores observados assumem o modelo:

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij},$$

em que o parâmetro de tratamento μ significa a média geral do experimento, sob a restrição $\sum_i \alpha_i = 0$. Dessa forma, a imagem de \mathbf{X} corresponde à decomposição de V_T em soma direta de $V_0 \oplus W_T$. E a imagem do eixo relativo ao parâmetro μ é justamente o subespaço V_0 .

Vale destacar que se compararmos a parametrização anterior com a atual, no modelo temos que $\tau_i = \mu + \alpha_i$.

4.2 Modelos lineares na presença de efeitos de blocos

O modelo para o delineamento em blocos casualizados, na notação matricial, considera a matriz de delineamento $\mathbf{X}_{N \times (t+b)}$ e o vetor de parâmetros $\boldsymbol{\beta}_{(t+b)}$ definidos por:

$$\mathbf{X} = \begin{pmatrix} \mathbf{X}_1 & | & \mathbf{X}_2 \end{pmatrix} \text{ e } \boldsymbol{\beta} = \begin{pmatrix} \tau_1 \\ \vdots \\ \tau_t \\ \zeta_1 \\ \vdots \\ \zeta_b \end{pmatrix},$$

sendo $\mathbf{X}_1 = (\mathbf{u}_1 \dots \mathbf{u}_t)$, $\mathbf{X}_2 = (\mathbf{v}_1 \dots \mathbf{v}_b)$, τ_1, \dots, τ_t as médias dos tratamentos e ζ_1, \dots, ζ_b as médias dos blocos. De modo que as parcelas são enumeradas pelos tratamentos e pelos blocos. Note que tanto \mathbf{X}_1 quanto \mathbf{X}_2 não possuem o vetor coluna correspondente a \mathbf{u}_0 .

Matricialmente, o modelo é expresso por:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon},$$

em que \mathbf{Y} é o vetor de dados, \mathbf{X} é a matriz de delineamento, $\boldsymbol{\beta}$ é o vetor de parâmetros e $\boldsymbol{\varepsilon}$ é o vetor de resíduos.

Note que $\boldsymbol{\beta}$ não é o mesmo que $\boldsymbol{\tau} + \boldsymbol{\zeta}$, definido na Seção 3.2, isso porque $\boldsymbol{\beta}_{(t+b) \times 1}$ enquanto $(\boldsymbol{\tau} + \boldsymbol{\zeta})_{N \times 1}$. De modo que $\boldsymbol{\tau} + \boldsymbol{\zeta} = \mathbf{X}\boldsymbol{\beta}$, ou seja, $\boldsymbol{\tau}$ contém as coordenadas relativas as repetições dos tratamentos e $\boldsymbol{\zeta}$ contém as coordenadas relativas as repetições dos blocos.

Na abordagem geométrica, novamente, \mathbf{X} é considerada uma transformação linear do espaço de parâmetros \mathbb{R}^{t+b} no espaço de dados \mathbb{R}^N . Na matriz \mathbf{X} temos que \mathbf{X}_1 é uma matriz $N \times t$, que corresponde ao subespaço de tratamento V_T e \mathbf{X}_2 é uma matriz $N \times b$, que corresponde ao subespaço de bloco V_B .

Note que toda linha de \mathbf{X} possui dois "1" indicando o tratamento e o bloco. Como a soma das colunas de \mathbf{X}_1 resulta no vetor \mathbf{u}_0 e o mesmo ocorre para a soma das colunas de \mathbf{X}_2 , temos que $V_T \cap V_B$ contém o subespaço, \mathbf{u}_0 . De modo que \mathbf{X} não tem posto completo e

$$\begin{aligned} \mathbf{X}^T \mathbf{X} &= \begin{pmatrix} \mathbf{X}_1^T & | & \mathbf{X}_2^T \end{pmatrix} \begin{pmatrix} \mathbf{X}_1 & | & \mathbf{X}_2 \end{pmatrix} \\ &= \begin{pmatrix} \mathbf{X}_1^T \mathbf{X}_1 & \mathbf{X}_1^T \mathbf{X}_2 \\ \mathbf{X}_2^T \mathbf{X}_1 & \mathbf{X}_2^T \mathbf{X}_2 \end{pmatrix}, \end{aligned}$$

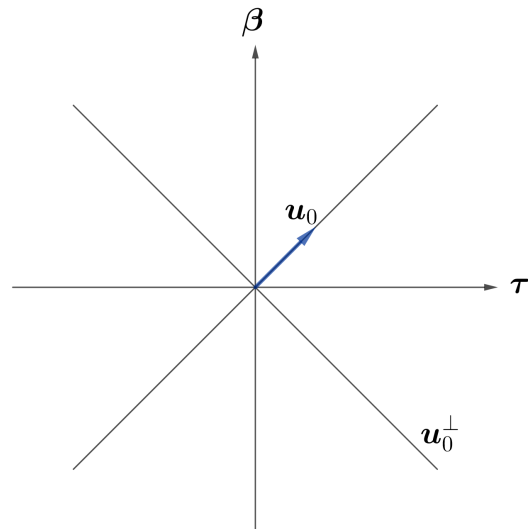
não possui inversa simples.

Para que $\mathbf{X} = (\mathbf{X}_1 | \mathbf{X}_2)$ seja injetiva é necessário restringir \mathbf{X} ao subespaço de \mathbb{R}^{t+b} ortogonal ao subespaço gerado por \mathbf{u}_0 (Figura 4.2).

A solução para o modelo matricial, pelo método dos mínimos quadrados, consiste em obter vetores em $Im(\mathbf{X}_1)$ e $Im(\mathbf{X}_2)$ o mais próximo possível do vetor de dados \mathbf{Y} . Novamente, isso é obtido pela projeção ortogonal de \mathbf{Y} em $Im(\mathbf{X}_1)$ e $Im(\mathbf{X}_2)$, ou seja, $P_{Im(\mathbf{X}_1)}\mathbf{Y}$ e $P_{Im(\mathbf{X}_2)}\mathbf{Y}$. Considerando \mathbf{X} restrita ao subespaço gerado por \mathbf{u}_0 , \mathbf{X} é injetiva, e a estimativa de mínimos quadrados é obtida pela pré-imagem de \mathbf{X} , isto é:

$$\mathbf{X}\hat{\boldsymbol{\beta}} = P_{Im(\mathbf{X})}\mathbf{Y} = P_{Im(\mathbf{X}_1)}\mathbf{Y} + P_{Im(\mathbf{X}_2)}\mathbf{Y}.$$

Figura 4.2 – Subespaço \mathbb{R}^{t+b} restrito ao subespaço gerado por \mathbf{u}_0 .



Fonte: Da autora (2019).

Note que $P_{Im(\mathbf{X}_1)}\mathbf{Y}$ é o vetor de médias ajustadas dos tratamentos e $P_{Im(\mathbf{X}_2)}\mathbf{Y}$ é o vetor de médias ajustadas dos blocos.

Na análise de variância, para o delineamento em blocos casualizados, assumimos que os valores observados assumem o modelo:

$$Y_{ij} = \tau_i + \zeta_j + \varepsilon_{ij},$$

para $i = 1, \dots, t$ (t número de tratamentos) e $j = 1, \dots, b$ (b número de blocos).

Quando se utilizam blocos, o espaço de parâmetros é dado por \mathbb{R}^{t+b} , de modo que os efeitos dos blocos são mensurados por um novo conjunto de parâmetros ζ_j , $j = 1, \dots, b$.

5 ANÁLISE DE VARIÂNCIA NA PRESENÇA DE EFEITOS ALEATÓRIOS

Em análise de variância, para modelos que apenas o resíduo é efeito aleatório, a matriz de variâncias e covariâncias Σ pode assumir diferentes estruturas, além de $\sigma^2 I$. Pois, mesmo considerando-se que apenas o resíduo é efeito aleatório pode existir correlação entre pares de parcelas distintas.

Por outro lado, é muito comum considerar modelos que não apenas o resíduo é efeito aleatório. Esse modelo é mais geral, e considera-se que outros efeitos sejam aleatórios. Para esse caso, a matriz de variâncias e covariância é dada por uma matriz simétrica Σ . Nesse modelo, as suposições sobre qual subespaço pertence o valor esperado de \mathbf{Y} não são alteradas.

Efeitos aleatórios são aqueles representativos de tratamentos de uma amostra proveniente de uma determinada população (BARBIN, 1993). Se de uma população retiramos uma amostra, as conclusões sobre essa amostra podem ser estendidas à população.

5.1 A métrica de Mahalanobis

Para modelos que não apenas o resíduo é efeito aleatório, a matriz de variâncias e covariâncias não mais é dada por $\sigma^2 I$, mas sim, por uma matriz simétrica Σ . Nesse caso, temos uma extensão do Teorema 3.1.3.

Teorema 5.1.1. *Assuma que $E(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$ e $cov(\mathbf{Y}) = \Sigma$. Sejam \mathbf{x} e \mathbf{z} quaisquer vetores em V_T . Então*

- i) *A variância do estimador $\mathbf{x} \cdot \mathbf{Y}$ é $\mathbf{x}^T \Sigma \mathbf{x}$;*
- ii) *A covariância de $\mathbf{x} \cdot \mathbf{Y}$ e $\mathbf{z} \cdot \mathbf{Y}$ é $\mathbf{x}^T \Sigma \mathbf{z}$;*
- iii) *Se \mathbf{x} é um autovetor¹ de Σ relativo ao autovalor δ , então a variância de $\mathbf{x} \cdot \mathbf{Y}$ é $\|\mathbf{x}\|^2 \delta$;*
- iv) *Sejam \mathbf{x} e \mathbf{z} autovetores de Σ correspondentes aos autovalores δ e η , respectivamente.*
 - a) *Se $\delta = \eta$, então $cov(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = (\mathbf{x} \cdot \mathbf{z})\eta$,*
 - b) *Se $\delta \neq \eta$, então $cov(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = 0$.*

Demonstração.

¹ Veja Definição .0.9 do Apêndice deste trabalho.

- i) Usando o fato de que para uma variável aleatória Y vale que $var(Y) = E((Y - E(Y))^2)$ temos que:

$$\begin{aligned}
 var(\mathbf{x} \cdot \mathbf{Y}) &= var(\mathbf{x}^\top \mathbf{Y}) \\
 &= E((\mathbf{x}^\top \mathbf{Y} - E(\mathbf{x}^\top \mathbf{Y}))^2) \\
 &= E((\mathbf{x}^\top \mathbf{Y} - E(\mathbf{x}^\top \mathbf{Y}))(\mathbf{x}^\top \mathbf{Y} - E(\mathbf{x}^\top \mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) \mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top \mathbf{x}) \\
 &= \mathbf{x}^\top E((\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top) \mathbf{x}.
 \end{aligned}$$

Como para variáveis aleatórias X e Y , vale que $cov(X, Y) = E((X - E(X))(Y - E(Y)))$ segue que:

$$\begin{aligned}
 var(\mathbf{x} \cdot \mathbf{Y}) &= \mathbf{x}^\top cov(\mathbf{Y}) \mathbf{x} \\
 &= \mathbf{x}^\top \Sigma \mathbf{x}.
 \end{aligned}$$

- ii) Usando o fato mencionado no item anterior que $cov(X, Y) = E((X - E(X))(Y - E(Y)))$, temos que:

$$\begin{aligned}
 cov(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) &= cov(\mathbf{x}^\top \mathbf{Y}, \mathbf{z}^\top \mathbf{Y}) \\
 &= E((\mathbf{x}^\top \mathbf{Y} - E(\mathbf{x}^\top \mathbf{Y}))(\mathbf{z}^\top \mathbf{Y} - E(\mathbf{z}^\top \mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) \mathbf{z}^\top (\mathbf{Y} - E(\mathbf{Y}))) \\
 &= E(\mathbf{x}^\top (\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top \mathbf{z}) \\
 &= \mathbf{x}^\top E((\mathbf{Y} - E(\mathbf{Y})) (\mathbf{Y} - E(\mathbf{Y}))^\top) \mathbf{z} \\
 &= \mathbf{x}^\top cov(\mathbf{Y}) \mathbf{z} \\
 &= \mathbf{x}^\top \Sigma \mathbf{z}.
 \end{aligned}$$

- iii) Pelo item i) temos que:

$$var(\mathbf{x} \cdot \mathbf{Y}) = \mathbf{x}^\top \Sigma \mathbf{x} = \mathbf{x}^\top (\delta \mathbf{x}) = (\mathbf{x}^\top \mathbf{x}) \delta = (\mathbf{x} \cdot \mathbf{x}) \delta = \|\mathbf{x}\|^2 \delta.$$

iv) a) Pelo item ii) temos que:

$$\text{cov}(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = \mathbf{x}^\top \boldsymbol{\Sigma} \mathbf{z} = \mathbf{x}^\top (\boldsymbol{\eta} \mathbf{z}) = (\mathbf{x}^\top \mathbf{z}) \boldsymbol{\eta} = (\mathbf{x} \cdot \mathbf{z}) \boldsymbol{\eta}.$$

b) Basta mostrar que sendo $\boldsymbol{\Sigma}$ uma matriz simétrica, autovetores \mathbf{x} e \mathbf{z} correspondentes a autovalores distintos são ortogonais entre si. Vejamos:

$$\begin{aligned} \delta(\mathbf{x} \cdot \mathbf{z}) &= (\delta \mathbf{x}) \cdot \mathbf{z} = \boldsymbol{\Sigma} \mathbf{x} \cdot \mathbf{z} = (\boldsymbol{\Sigma} \mathbf{x})^\top \mathbf{z} \\ &= (\mathbf{x}^\top \boldsymbol{\Sigma}^\top) \mathbf{z} = (\mathbf{x}^\top \boldsymbol{\Sigma}) \cdot \mathbf{z} = \mathbf{x}^\top (\boldsymbol{\Sigma} \mathbf{z}) \\ &= \mathbf{x}^\top (\boldsymbol{\eta} \mathbf{z}) = \boldsymbol{\eta} (\mathbf{x}^\top \mathbf{z}) = \boldsymbol{\eta} (\mathbf{x} \cdot \mathbf{z}). \end{aligned}$$

Logo, $(\delta - \boldsymbol{\eta})(\mathbf{x} \cdot \mathbf{z}) = 0$, mas como $\delta \neq \boldsymbol{\eta}$, então $\mathbf{x} \cdot \mathbf{z} = 0$. Portanto,

$$\text{cov}(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = (\mathbf{x} \cdot \mathbf{z}) \boldsymbol{\eta} = 0.$$

□

Como a matriz de variâncias e covariâncias é dada por uma matriz simétrica $\boldsymbol{\Sigma}$, vimos no Teorema 5.1.1 que $\text{cov}(\mathbf{x} \cdot \mathbf{Y}, \mathbf{z} \cdot \mathbf{Y}) = \mathbf{x}^\top \boldsymbol{\Sigma} \mathbf{z}$. Em razão desse fato, a métrica usual para vetores em \mathbb{R}^N é métrica

$$\langle\langle \mathbf{x}, \mathbf{z} \rangle\rangle = \mathbf{x}^\top \boldsymbol{\Sigma} \mathbf{z},$$

que é denominada métrica de Mahalanobis.

Para o caso em que a matriz de variâncias e covariâncias é dada por $\sigma^2 \mathbf{I}$, temos que $\boldsymbol{\Sigma} = \sigma^2 \mathbf{I}$, logo, $\mathbf{x}^\top \sigma^2 \mathbf{I} \mathbf{z} = (\mathbf{x} \cdot \mathbf{z}) \sigma^2$. Portanto, a métrica de Mahalanobis se torna a métrica Euclidiana usual.

Mesmo na presença de efeitos aleatórios, em análise de variância estamos interessados em realizar um teste. Vimos na Seção 3.1 que os pressupostos da análise de variância são: aditividade, normalidade, independência e homogeneidade de variâncias.

Na presença de efeitos aleatórios a suposição de independência fica comprometida. Contudo, na presença de covariância, tais experimentos podem se analisados. Isso porque, o teste de hipótese é realizado considerando autovetores relativos ao mesmo autovalor, como mostra o teorema a seguir.

Teorema 5.1.2. *Suponha que \mathbf{Y} tem distribuição normal multivariada, $E(\mathbf{Y}) = \boldsymbol{\tau} \in V_T$ e $\text{cov}(\mathbf{Y}) = \boldsymbol{\Sigma}$. Então temos o seguinte.*

- i) *Seja W_1 um subespaço vetorial d_1 -dimensional, consistindo de autovetores de $\boldsymbol{\Sigma}$ relativos ao autovalor δ , sendo W_1 ortogonal a V_T . Então $\|P_{W_1}\mathbf{Y}\|^2/\delta$ tem distribuição χ^2 com d_1 graus de liberdade.*
- ii) *Sejam W_1 e W_2 subespaços vetoriais com dimensões d_1 e d_2 respectivamente, ambos consistindo de autovetores de $\boldsymbol{\Sigma}$ com autovalor δ , sendo W_1 e W_2 ortogonais entre si. Então $QM(W_1)/QM(W_2)$ tem distribuição F com d_1 e d_2 graus de liberdade.*

Demonstração.

- i) *Seja $\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_{d_1}$ uma base ortonormal para W_1 . Expressando $P_{W_1}\mathbf{Y}$ com relação a essa base temos:*

$$\begin{aligned} P_{W_1}\mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{u}_1}{\mathbf{u}_1 \cdot \mathbf{u}_1} \right) \mathbf{u}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_2}{\mathbf{u}_2 \cdot \mathbf{u}_2} \right) \mathbf{u}_2 + \dots + \left(\frac{\mathbf{Y} \cdot \mathbf{u}_{d_1}}{\mathbf{u}_{d_1} \cdot \mathbf{u}_{d_1}} \right) \mathbf{u}_{d_1} \\ &= (\mathbf{Y} \cdot \mathbf{u}_1) \mathbf{u}_1 + (\mathbf{Y} \cdot \mathbf{u}_2) \mathbf{u}_2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1}) \mathbf{u}_{d_1} \end{aligned}$$

Calculando a esperança de $\mathbf{Y} \cdot \mathbf{u}_i$ e usando o fato que W_1 é ortogonal a V_T temos:

$$E(\mathbf{Y} \cdot \mathbf{u}_1) = E(\mathbf{Y}) \cdot \mathbf{u}_1 = \boldsymbol{\tau} \cdot \mathbf{u}_1 = 0.$$

Calculando a variância de $\mathbf{Y} \cdot \mathbf{u}_i$ temos:

$$\text{var}(\mathbf{Y} \cdot \mathbf{u}_i) = \mathbf{u}_i^T \boldsymbol{\Sigma} \mathbf{u}_i = \|\mathbf{u}_i\|^2 \delta = \delta.$$

Logo, $\mathbf{Y} \cdot \mathbf{u}_i$ são variáveis normais com média 0 e variância δ . Portanto segue que:

$$\begin{aligned} \frac{\|P_{W_1}\mathbf{Y}\|^2}{\delta} &= \frac{1}{\delta} ((\mathbf{Y} \cdot \mathbf{u}_1)^2 + (\mathbf{Y} \cdot \mathbf{u}_2)^2 + \dots + (\mathbf{Y} \cdot \mathbf{u}_{d_1})^2) \\ &= \frac{1}{\delta} \left(\sum_{i=1}^{d_1} (\mathbf{Y} \cdot \mathbf{u}_i)^2 \right) \\ &= \sum_{i=1}^{d_1} \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sqrt{\delta}} \right)^2. \end{aligned}$$

Como $\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sqrt{\delta}}$ tem distribuição normal padrão, segue que $\sum_{i=1}^{d_1} \left(\frac{\mathbf{Y} \cdot \mathbf{u}_i}{\sqrt{\delta}}\right)^2$ tem distribuição χ^2 com d_1 graus de liberdade.

ii) Como W_1 e W_2 são subespaços ortogonais, $P_{W_1}\mathbf{Y}$ e $P_{W_2}\mathbf{Y}$ são vetores aleatórios independentes. Como também $P_{W_1}\mathbf{Y}$ e $P_{W_2}\mathbf{Y}$ são vetores relativos ao mesmo autovalor δ segue que:

$$\frac{\|P_{W_1}\mathbf{Y}\|^2}{\delta} \text{ e } \frac{\|P_{W_2}\mathbf{Y}\|^2}{\delta}$$

têm distribuição χ^2 com d_1 e d_2 graus de liberdade respectivamente. Logo, segue que

$$\frac{\|P_{W_1}\mathbf{Y}\|^2/\delta}{\|P_{W_2}\mathbf{Y}\|^2/\delta} = \frac{\|P_{W_1}\mathbf{Y}\|^2}{\|P_{W_2}\mathbf{Y}\|^2}$$

tem distribuição F com d_1 e d_2 graus de liberdade. Portanto,

$$\frac{\|P_{W_1}\mathbf{Y}\|^2/d_1}{\|P_{W_2}\mathbf{Y}\|^2/d_2} = \frac{QM(W_1)}{QM(W_2)}$$

tem de distribuição F .

□

O teorema anterior descreve o caso em que os autovetores de Σ correspondem ao mesmo autovalor. Nesse caso, a razão de variâncias é sempre entre χ^2 independentes. E como os autovalores são os mesmos essa razão possui distribuição F . Desse modo, as suposições de independência e homogeneidade de variâncias, sob as condições do teorema, não são afetadas.

Para o caso que os autovetores de Σ correspondem a autovetores distintos, vimos no Teorema 5.1.1, item iv), item b) que esses são sempre ortogonais. Assim, se \mathbf{x} e \mathbf{z} são autovetores de Σ com autovalores ξ_0 e ξ_1 . Como Σ é uma matriz simétrica temos que $\mathbf{x} \cdot \mathbf{z} = 0$, ou seja, autovetores, associados a autovalores distintos, de matrizes simétricas são sempre ortogonais². Desse modo, a suposição de independência é válida.

De modo geral, para realizar o teste F , na presença de efeitos aleatórios, consideramos os autoespaços³ da matriz Σ .

² Esse é um resultado de álgebra linear que encontra-se no Apêndice deste trabalho, Teorema .0.2

³ O conjunto de todos os vetores associados a um mesmo autovalor é denominado autoespaço correspondente a esse autovalor.

5.2 Análise de variância na presença de covariância

Nas próximas seções, apresentamos a análise de variâncias na presença de efeitos aleatórios para os delineamentos apresentados nas Seções 3.1, 3.2 e 3.3.

Para a análise de variância, os subespaços de interesse são: $V, V_0, V_T, W_T, V_B, W_B, V_L, W_L, V_C, W_C, W_E$, dentre outros. A análise de variância para delineamentos particulares será possível, se os subespaços de interesse forem ortogonais, na métrica de Mahalanobis, o que implica não correlação e, na suposição de normalidade, independência.

5.2.1 Efeitos aleatórios no delineamento inteiramente casualizado

Na ausência de efeitos de blocos, na análise de variância é razoável assumir um modelo onde a covariância \mathbf{Y} para parcelas α e β é dada por:

$$\text{cov}(Y_\alpha, Y_\beta) = \begin{cases} \sigma^2 & \text{se } \alpha = \beta \\ \rho\sigma^2 & \text{se } \alpha \neq \beta. \end{cases}$$

Portanto, a matriz de variâncias e covariâncias é dada por:

$$\begin{aligned} \mathbf{\Sigma} &= \text{cov}(\mathbf{Y}) \\ &= \sigma^2 \mathbf{I} + \rho\sigma^2(\mathbf{J} - \mathbf{I}) \\ &= \sigma^2[\mathbf{I} + \rho(\mathbf{J} - \mathbf{I})] \\ &= \sigma^2 \begin{pmatrix} 1 & \rho & \dots & \rho \\ \rho & 1 & \dots & \rho \\ \dots & \dots & \dots & \dots \\ \rho & \rho & \dots & 1 \end{pmatrix}, \end{aligned}$$

em que \mathbf{I} é a matriz identidade e \mathbf{J} é uma matriz $N \times N$ de uns.

Para o delineamento inteiramente casualizado, agora decomponos V em soma direta de V_0 e V_0^\perp , e para $\mathbf{u}_0 \in V_0$ e $\mathbf{v} \in V_0^\perp$ obtemos:

$$\begin{aligned}\boldsymbol{\Sigma}\mathbf{u}_0 &= \sigma^2[\mathbf{I} + \rho(\mathbf{J} - \mathbf{I})]\mathbf{u}_0 \\ &= \sigma^2[\mathbf{I}\mathbf{u}_0 + \rho(\mathbf{J}\mathbf{u}_0 - \mathbf{I}\mathbf{u}_0)] \\ &= \sigma^2[\mathbf{u}_0 + \rho(N\mathbf{u}_0 - \mathbf{u}_0)] \\ &= \sigma^2[1 + \rho(N - 1)]\mathbf{u}_0 \\ &= \sigma^2(1 + N\rho - \rho)\mathbf{u}_0,\end{aligned}$$

e portanto, \mathbf{u}_0 é autovetor relativo ao autovalor $\xi_0 = \sigma^2(1 + N\rho - \rho)$.

Agora note que $\mathbf{v} = (v_1, \dots, v_N)$ é tal que $\mathbf{v} \cdot \mathbf{u}_0 = 0$, então $\sum_i v_i = 0$ e isto implica que $\mathbf{J}\mathbf{v} = \mathbf{0}$. Assim temos que:

$$\begin{aligned}\boldsymbol{\Sigma}\mathbf{v} &= \sigma^2[\mathbf{I} + \rho(\mathbf{J} - \mathbf{I})]\mathbf{v} \\ &= \sigma^2[\mathbf{I}\mathbf{v} + \rho(\mathbf{J}\mathbf{v} - \mathbf{I}\mathbf{v})] \\ &= \sigma^2[\mathbf{v} + \rho(\mathbf{0} - \mathbf{v})] \\ &= \sigma^2[1 + \rho(0 - 1)]\mathbf{v} \\ &= \sigma^2(1 - \rho)\mathbf{v},\end{aligned}$$

e portanto, \mathbf{v} é um autovetor relativo ao autovalor $\xi_1 = \sigma^2(1 - \rho)$.

Observe que $\dim(V_0) = 1$ e $\dim(V_0^\perp) = N - 1$, logo, por dimensão foram obtidos todos os autoespaços de $\boldsymbol{\Sigma}$.

A análise de variância é feita como antes, porém, agora considerando outovetores relativos aos mesmos autovalores. Como vimos na Seção 3.1, para o delinemenro inteiramente casualizado V é decomposto em soma direta de $V_0 \oplus W_T \oplus V_T^\perp$. De modo que o autovalor ξ_0 corresponde ao autoespaço V_0 e o autovalor ξ_1 corresponde aos autoespaços W_T e V_T^\perp . A razão de variâncias é calculada em termos dos mesmo autovalores, veja Tabela 5.1.

Vale destacar que ξ_0 não pode ser estimado, de modo que não é possível verificar se a média geral do experimento é estatisticamente diferente de zero. Entretanto, o teste para as médias dos tratamentos pode ser realizado normalmente, pois ξ_1 pode ser estimado por $QM(V_T^\perp) = \hat{\xi}_1$.

Tabela 5.1 – Tabela de análise de variância para delineamento inteiramente casualizado na presença de efeitos aleatórios.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	
Tratamentos W_T	$t - 1$	$\ P_{W_T}\mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/\hat{\xi}_1$
Resíduo V_T^\perp	$N - t$	$\ P_{V_T^\perp}\mathbf{Y}\ ^2$	$QM(V_T^\perp) = \hat{\xi}_1$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

5.2.2 Efeitos aleatórios no delineamento em blocos casualizados

Outro caso que pode ser analisado é a presença de efeito de blocos no modelo de análise de variância. No Tópico 3.2 os blocos foram considerados como efeitos fixos, em que a matriz de variâncias e covariâncias era $\sigma^2\mathbf{I}$. Contudo, é comum considerar um modelo mais geral, em que os efeitos dos blocos sejam aleatórios. A covariância de \mathbf{Y} para parcelas α e β fica:

$$cov(Y_\alpha, Y_\beta) = \begin{cases} \sigma^2 & \text{se } \alpha = \beta \\ \rho_1 \sigma^2 & \text{se } \alpha \neq \beta \text{ mas } \alpha \text{ e } \beta \text{ estão no mesmo bloco} \\ \rho_2 \sigma^2 & \text{se } \alpha \neq \beta \text{ mas } \alpha \text{ e } \beta \text{ não estão no mesmo bloco.} \end{cases}$$

A covariância entre respostas de pares de parcelas depende se estão ou não em blocos distintos. Em geral, se supõe que a covariância entre respostas de parcelas que estão em blocos distintos é maior do que a covariância entre parcelas que estão no mesmo bloco. Assim, a matriz

de variâncias e covariâncias é:

$$\begin{aligned}
\boldsymbol{\Sigma} &= \text{cov}(\mathbf{Y}) \\
&= \sigma^2 \mathbf{I} + \rho_1 \sigma^2 (\mathbf{J}_B - \mathbf{I}) + \rho_2 \sigma^2 (\mathbf{J} - \mathbf{J}_B) \\
&= \sigma^2 [\mathbf{I} + \rho_1 (\mathbf{J}_B - \mathbf{I}) + \rho_2 (\mathbf{J} - \mathbf{J}_B)] \\
&= \sigma^2 \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 & \dots & \rho_2 & \rho_2 & \dots & \rho_2 \\ \rho_1 & 1 & \dots & \rho_1 & \dots & \rho_2 & \rho_2 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_1 & \rho_1 & \dots & 1 & \dots & \rho_2 & \rho_2 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_2 & \rho_2 & \dots & \rho_2 & \dots & 1 & \rho_1 & \dots & \rho_1 \\ \rho_2 & \rho_2 & \dots & \rho_2 & \dots & \rho_1 & 1 & \dots & \rho_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_2 & \rho_2 & \dots & \rho_2 & \dots & \rho_1 & \rho_1 & \dots & 1 \end{pmatrix},
\end{aligned}$$

em que \mathbf{I} é a matriz identidade, \mathbf{J} é uma matriz $N \times N$ de uns e \mathbf{J}_B é uma matriz $N \times N$ de blocos.

Para exemplificar, seja um experimento com dois blocos, ambos de tamanho $k = 3$, as matrizes \mathbf{J}_B e $\boldsymbol{\Sigma}$ são dadas respectivamente por:

$$\mathbf{J}_B = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix} \text{ e } \boldsymbol{\Sigma} = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_1 & \rho_2 & \rho_2 & \rho_2 \\ \rho_1 & 1 & \rho_1 & \rho_2 & \rho_2 & \rho_2 \\ \rho_1 & \rho_1 & 1 & \rho_2 & \rho_2 & \rho_2 \\ \rho_2 & \rho_2 & \rho_2 & 1 & \rho_1 & \rho_1 \\ \rho_2 & \rho_2 & \rho_2 & \rho_1 & 1 & \rho_1 \\ \rho_2 & \rho_2 & \rho_2 & \rho_1 & \rho_1 & 1 \end{pmatrix}.$$

Para o delineamento em blocos, temos que $V_B = V_0 \oplus W_B$, desse modo $V = V_B \oplus V_B^\perp = V_0 \oplus W_B \oplus V_B^\perp$. Sejam \mathbf{u}_0 , \mathbf{x} e \mathbf{z} vetores de V_0 , W_B e V_B^\perp , respectivamente.

$$\begin{aligned}
\Sigma \mathbf{u}_0 &= \sigma^2 [\mathbf{I} + \rho_1 (\mathbf{J}_B - \mathbf{I}) + \rho_2 (\mathbf{J} - \mathbf{J}_B)] \mathbf{u}_0 \\
&= \sigma^2 [\mathbf{I} \mathbf{u}_0 + \rho_1 (\mathbf{J}_B \mathbf{u}_0 - \mathbf{I} \mathbf{u}_0) + \rho_2 (\mathbf{J} \mathbf{u}_0 - \mathbf{J}_B \mathbf{u}_0)] \\
&= \sigma^2 [\mathbf{u}_0 - \rho_1 (k \mathbf{u}_0 - \mathbf{u}_0) + \rho_2 (N \mathbf{u}_0 - k \mathbf{u}_0)] \\
&= \sigma^2 [1 - \rho_1 (k - 1) + \rho_2 (N - k)] \mathbf{u}_0,
\end{aligned}$$

$$\begin{aligned}
\Sigma \mathbf{x} &= \sigma^2 [\mathbf{I} + \rho_1 (\mathbf{J}_B - \mathbf{I}) + \rho_2 (\mathbf{J} - \mathbf{J}_B)] \mathbf{x} \\
&= \sigma^2 [\mathbf{I} \mathbf{x} + \rho_1 (\mathbf{J}_B \mathbf{x} - \mathbf{I} \mathbf{x}) + \rho_2 (\mathbf{J} \mathbf{x} - \mathbf{J}_B \mathbf{x})] \\
&= \sigma^2 [\mathbf{x} + \rho_1 (k \mathbf{x} - \mathbf{x}) + \rho_2 (\mathbf{0} - k \mathbf{x})] \\
&= \sigma^2 [1 + \rho_1 (k - 1) + \rho_2 (0 - k)] \mathbf{x} \\
&= \sigma^2 [1 + \rho_1 (k - 1) - k \rho_2] \mathbf{x},
\end{aligned}$$

$$\begin{aligned}
\Sigma \mathbf{z} &= \sigma^2 [\mathbf{I} + \rho_1 (\mathbf{J}_B - \mathbf{I}) + \rho_2 (\mathbf{J} - \mathbf{J}_B)] \mathbf{z} \\
&= \sigma^2 [\mathbf{I} \mathbf{z} + \rho_1 (\mathbf{J}_B \mathbf{z} - \mathbf{I} \mathbf{z}) + \rho_2 (\mathbf{J} \mathbf{z} - \mathbf{J}_B \mathbf{z})] \\
&= \sigma^2 [\mathbf{z} + \rho_1 (\mathbf{0} - \mathbf{z}) + \rho_2 (\mathbf{0} - \mathbf{0})] \\
&= \sigma^2 [1 + \rho_1 (-1) + \rho_2 (0 - 0)] \mathbf{z} \\
&= \sigma^2 (1 - \rho_1) \mathbf{z}.
\end{aligned}$$

Portanto, $\xi_0 = \sigma^2 [1 - \rho_1 (k - 1) + \rho_2 (N - k)]$, $\xi_1 = \sigma^2 [1 - \rho_1 (k - 1) + k \rho_2]$ e $\xi_2 = \sigma^2 (1 - \rho_1)$ são os autovalores correspondentes aos autovetores \mathbf{u}_0 , \mathbf{x} e \mathbf{z} respectivamente.

Como $\dim(V_0) = 1$, $\dim(W_B) = b - 1$ e $\dim(V_B^\perp) = N - b$, por dimensão, estes são todos os autoespaços de Σ . Na Seção 3.2 para o delineamento em blocos vimos que o espaço vetorial V pode ser decomposto em soma direta de:

$$\begin{aligned}
V &= (V_T + V_B) \oplus (V_T + V_B)^\perp \\
&= V_0 \oplus W_B \oplus W_T \oplus (V_T + V_B)^\perp \\
&= V_0 \oplus W_B \oplus W_T \oplus W_E.
\end{aligned}$$

Assim, o subespaço W_T está contido no autoespaço V_B^\perp e o autovalor ξ_2 corresponde aos subespaços W_T e W_E . A análise de variância é obtida considerando outovetores correspondentes aos mesmos autovalores, como mostra a Tabela 5.2. Assim, a razão de variâncias é calculada em termos dos mesmos autovalores.

Tabela 5.2 – Tabela de análise de variância para delineamento em blocos casualizados na presença de efeitos aleatórios.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	
Blocos W_B	$b - 1$	$\ P_{W_B}\mathbf{Y}\ ^2$	$QM(W_B) = \hat{\xi}_1$	
Tratamentos W_T	$t - 1$	$\ P_{W_T}\mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/\hat{\xi}_2$
Resíduo W_E	$N - b - t + 1$	$\ P_{W_E}\mathbf{Y}\ ^2$	$QM(W_E) = \hat{\xi}_2$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

Novamente, ξ_0 não pode ser estimado, o que não prejudica o teste de hipótese para tratamentos, uma vez que este é realizado pela estimativa de ξ_2 .

Se os blocos produzem efeitos aleatórios a hipótese testada refere-se a variabilidade dos blocos. Para realizar o teste basta comparar $\hat{\xi}_1/\hat{\xi}_2$. Se esse valor é maior que um o teste é significativo. Logo, existe variabilidade entre os blocos. As conclusões obtidas para este teste podem estendidas a toda a população, ou seja, há diferença significativa entre os blocos de toda população.

5.2.3 Efeitos aleatórios no delineamento em linhas e colunas

Como destaca Bailey (2008), também é possível considerar as linhas e as colunas como efeitos aleatórios. A covariância de \mathbf{Y} para parcelas α e β fica:

$$cov(Y_\alpha, Y_\beta) = \begin{cases} \sigma^2 & \text{se } \alpha = \beta \\ \rho_1 \sigma^2 & \text{se } \alpha \neq \beta \text{ mas } \alpha \text{ e } \beta \text{ estão na mesma linha} \\ \rho_2 \sigma^2 & \text{se } \alpha \neq \beta \text{ mas } \alpha \text{ e } \beta \text{ estão na mesma coluna} \\ \rho_3 \sigma^2 & \text{caso contrário.} \end{cases}$$

De modo que, a covariância entre pares de parcelas depende se estão na mesma linha e também na mesma coluna. A matriz de variâncias e covariâncias é dada por:

$$\begin{aligned}
 \boldsymbol{\Sigma} &= \text{cov}(\mathbf{Y}) \\
 &= \sigma^2 \mathbf{I} + \rho_1 \sigma^2 (\mathbf{J}_L - \mathbf{I}) + \rho_2 \sigma^2 (\mathbf{J}_C - \mathbf{I}) + \rho_3 \sigma^2 (\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I}) \\
 &= \sigma^2 [\mathbf{I} + \rho_1 (\mathbf{J}_L - \mathbf{I}) + \rho_2 (\mathbf{J}_C - \mathbf{I}) + \rho_3 (\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I})] \\
 &= \sigma^2 \begin{pmatrix} 1 & \rho_1 & \dots & \rho_1 & \dots & \rho_2 & \rho_3 & \dots & \rho_3 \\ \rho_1 & 1 & \dots & \rho_1 & \dots & \rho_3 & \rho_2 & \dots & \rho_3 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_1 & \rho_1 & \dots & 1 & \dots & \rho_3 & \rho_3 & \dots & \rho_2 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_2 & \rho_3 & \dots & \rho_3 & \dots & 1 & \rho_1 & \dots & \rho_1 \\ \rho_3 & \rho_2 & \dots & \rho_3 & \dots & \rho_1 & 1 & \dots & \rho_1 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \rho_3 & \rho_3 & \dots & \rho_2 & \dots & \rho_1 & \rho_1 & \dots & 1 \end{pmatrix}
 \end{aligned}$$

em que \mathbf{I} é a matriz identidade, \mathbf{J} é uma matriz $N \times N$ de uns, \mathbf{J}_L é uma matriz $N \times N$ de linhas e \mathbf{J}_C é uma matriz $N \times N$ de colunas.

Para exemplificar, seja um experimento com três linhas a três colunas, ou seja $m = n = 3$, as matrizes \mathbf{J}_L , \mathbf{J}_C e $\boldsymbol{\Sigma}$ são, respectivamente:

$$\mathbf{J}_L = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \quad \mathbf{J}_C = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix},$$

$$\Sigma = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_1 & \rho_2 & \rho_3 & \rho_3 & \rho_2 & \rho_3 & \rho_3 \\ \rho_1 & 1 & \rho_1 & \rho_3 & \rho_2 & \rho_3 & \rho_3 & \rho_2 & \rho_3 \\ \rho_1 & \rho_1 & 1 & \rho_3 & \rho_3 & \rho_2 & \rho_3 & \rho_3 & \rho_2 \\ \rho_2 & \rho_3 & \rho_3 & 1 & \rho_1 & \rho_1 & \rho_2 & \rho_3 & \rho_3 \\ \rho_3 & \rho_2 & \rho_3 & \rho_1 & 1 & \rho_1 & \rho_3 & \rho_2 & \rho_3 \\ \rho_3 & \rho_3 & \rho_2 & \rho_1 & \rho_1 & 1 & \rho_3 & \rho_3 & \rho_2 \\ \rho_2 & \rho_3 & \rho_3 & \rho_2 & \rho_3 & \rho_3 & 1 & \rho_1 & \rho_1 \\ \rho_3 & \rho_2 & \rho_3 & \rho_3 & \rho_2 & \rho_3 & \rho_1 & 1 & \rho_1 \\ \rho_3 & \rho_3 & \rho_2 & \rho_3 & \rho_3 & \rho_2 & \rho_1 & \rho_1 & 1 \end{pmatrix}.$$

Para o delineamento em linhas e colunas, temos que $V_L = V_0 \oplus W_L$ e $V_C = V_0 \oplus W_L$, de modo que $V = (V_L + V_C) \oplus (V_L + V_C)^\perp = V_0 \oplus W_L \oplus W_C \oplus (V_L + V_C)^\perp$. Sejam \mathbf{u}_0 , \mathbf{x} , \mathbf{y} e \mathbf{z} vetores em V_0 , W_L , W_C e $(V_L + V_C)^\perp$, respectivamente tem-se:

$$\begin{aligned} \Sigma \mathbf{u}_0 &= \sigma^2 [\mathbf{I} + \rho_1(\mathbf{J}_L - \mathbf{I}) + \rho_2(\mathbf{J}_C - \mathbf{I}) + \rho_3(\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I})] \mathbf{u}_0 \\ &= \sigma^2 [\mathbf{I} \mathbf{u}_0 + \rho_1(\mathbf{J}_L \mathbf{u}_0 - \mathbf{I} \mathbf{u}_0) + \rho_2(\mathbf{J}_C \mathbf{u}_0 - \mathbf{I} \mathbf{u}_0) + \rho_3(\mathbf{J} \mathbf{u}_0 - \mathbf{J}_L \mathbf{u}_0 - \mathbf{J}_C \mathbf{u}_0 + \mathbf{I} \mathbf{u}_0)] \\ &= \sigma^2 [\mathbf{u}_0 + \rho_1(n \mathbf{u}_0 - \mathbf{u}_0) + \rho_2(m \mathbf{u}_0 - \mathbf{u}_0) + \rho_3(N \mathbf{u}_0 - n \mathbf{u}_0 - m \mathbf{u}_0 + \mathbf{u}_0)] \\ &= \sigma^2 [1 + \rho_1(n - 1) + \rho_2(m - 1) + \rho_3(N - n - m + 1)] \mathbf{u}_0, \end{aligned}$$

$$\begin{aligned} \Sigma \mathbf{x} &= \sigma^2 [\mathbf{I} + \rho_1(\mathbf{J}_L - \mathbf{I}) + \rho_2(\mathbf{J}_C - \mathbf{I}) + \rho_3(\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I})] \mathbf{x} \\ &= \sigma^2 [\mathbf{I} \mathbf{x} + \rho_1(\mathbf{J}_L \mathbf{x} - \mathbf{I} \mathbf{x}) + \rho_2(\mathbf{J}_C \mathbf{x} - \mathbf{I} \mathbf{x}) + \rho_3(\mathbf{J} \mathbf{x} - \mathbf{J}_L \mathbf{x} - \mathbf{J}_C \mathbf{x} + \mathbf{I} \mathbf{x})] \\ &= \sigma^2 [\mathbf{x} + \rho_1(n \mathbf{x} - \mathbf{x}) + \rho_2(\mathbf{0} - \mathbf{x}) + \rho_3(\mathbf{0} - n \mathbf{x} - \mathbf{0} + \mathbf{x})] \\ &= \sigma^2 [1 + \rho_1(n - 1) + \rho_2(0 - 1) + \rho_3(0 - n - 0 + 1)] \mathbf{x} \\ &= \sigma^2 [1 + \rho_1(n - 1) - \rho_2 - \rho_3(n - 1)] \mathbf{x}, \end{aligned}$$

$$\begin{aligned}
\Sigma \mathbf{y} &= \sigma^2 [\mathbf{I} + \rho_1(\mathbf{J}_L - \mathbf{I}) + \rho_2(\mathbf{J}_C - \mathbf{I}) + \rho_3(\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I})] \mathbf{y} \\
&= \sigma^2 [\mathbf{I} \mathbf{y} + \rho_1(\mathbf{J}_L \mathbf{y} - \mathbf{I} \mathbf{y}) + \rho_2(\mathbf{J}_C \mathbf{y} - \mathbf{I} \mathbf{y}) + \rho_3(\mathbf{J} \mathbf{y} - \mathbf{J}_L \mathbf{y} - \mathbf{J}_C \mathbf{y} + \mathbf{I} \mathbf{y})] \\
&= \sigma^2 [\mathbf{y} + \rho_1(\mathbf{0} - \mathbf{y}) + \rho_2(m \mathbf{y} - \mathbf{y}) + \rho_3(\mathbf{0} - \mathbf{0} - m \mathbf{y} + \mathbf{y})] \\
&= \sigma^2 [1 + \rho_1(0 - 1) + \rho_2(m - 1) + \rho_3(0 - 0 - m + 1)] \mathbf{y} \\
&= \sigma^2 [1 - \rho_1 + \rho_2(m - 1) - \rho_3(m - 1)] \mathbf{y},
\end{aligned}$$

$$\begin{aligned}
\Sigma \mathbf{z} &= \sigma^2 [\mathbf{I} + \rho_1(\mathbf{J}_L - \mathbf{I}) + \rho_2(\mathbf{J}_C - \mathbf{I}) + \rho_3(\mathbf{J} - \mathbf{J}_L - \mathbf{J}_C + \mathbf{I})] \mathbf{z} \\
&= \sigma^2 [\mathbf{I} \mathbf{z} + \rho_1(\mathbf{J}_L \mathbf{z} - \mathbf{I} \mathbf{z}) + \rho_2(\mathbf{J}_C \mathbf{z} - \mathbf{I} \mathbf{z}) + \rho_3(\mathbf{J} \mathbf{z} - \mathbf{J}_L \mathbf{z} - \mathbf{J}_C \mathbf{z} + \mathbf{I} \mathbf{z})] \\
&= \sigma^2 [\mathbf{z} + \rho_1(\mathbf{0} - \mathbf{z}) + \rho_2(\mathbf{0} - \mathbf{z}) + \rho_3(\mathbf{0} - \mathbf{0} - \mathbf{0} + \mathbf{z})] \\
&= \sigma^2 [1 + \rho_1(0 - 1) + \rho_2(0 - 1) + \rho_3(0 - 0 - 0 + 1)] \mathbf{z} \\
&= \sigma^2 (1 - \rho_1 - \rho_2 + \rho_3) \mathbf{z}.
\end{aligned}$$

Portanto, $\xi_0 = \sigma^2 [1 + \rho_1(n - 1) + \rho_2(m - 1) + \rho_3(N - n - m + 1)]$, $\xi_1 = \sigma^2 [1 + \rho_1(n - 1) - \rho_2 - \rho_3(n - 1)]$, $\xi_2 = \sigma^2 [1 - \rho_1 + \rho_2(m - 1) - \rho_3(m - 1)]$ e $\xi_3 = \sigma^2 (1 - \rho_1 - \rho_2 + \rho_3)$ são autovalores correspondentes aos autovetores \mathbf{u}_0 , \mathbf{x} , \mathbf{y} e \mathbf{z} , respectivamente.

Novamente, por dimensão, os autoespaços de Σ são: V_0 , W_L , W_C e $(V_L + V_C)^\perp$. Vimos na Seção 3.3 que para o delineamento em linhas e colunas o espaço vetorial V pode ser decomposto em uma soma direta dada por:

$$\begin{aligned}
V &= (V_T + V_L + V_C) \oplus (V_T + V_L + V_C)^\perp \\
&= V_0 \oplus W_L \oplus W_C \oplus W_T (V_T + V_L + V_C)^\perp \\
&= V_0 \oplus W_L \oplus W_C \oplus W_T \oplus W_E.
\end{aligned}$$

Assim, o subespaço W_T está contido no autoespaço $(V_L + V_C)^\perp$ e o autovalor ξ_3 corresponde aos subespaços W_T e W_E . A análise de variância é realizada considerando autovetores relativos aos mesmos autovalores, como mostra a Tabela 5.3. A razão de variâncias é calculada em termos dos mesmos autovalores.

Se as linhas e as colunas produzem efeitos aleatórios as hipóteses testadas referem-se a variabilidade entre as linhas e entre as colunas. Para realizar o teste basta comparar $\hat{\xi}_1/\hat{\xi}_3$ e $\hat{\xi}_2/\hat{\xi}_3$. Se esses valores são maiores que um o teste é significativo. Logo, existe variabilidade entre as linhas e entre as colunas. As conclusões obtidas para este teste podem estendidas a

Tabela 5.3 – Tabela de análise de variância para delineamento em linhas e colunas na presença de efeitos aleatórios.

Fontes de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística F
Média V_0	1	$\ P_{V_0}\mathbf{Y}\ ^2$	$QM(V_0)$	
Linhas W_L	$m - 1$	$\ P_{W_L}\mathbf{Y}\ ^2$	$QM(W_L) = \hat{\xi}_1$	
Colunas W_C	$n - 1$	$\ P_{W_C}\mathbf{Y}\ ^2$	$QM(W_C) = \hat{\xi}_2$	
Tratamentos W_T	$t - 1$	$\ P_{W_T}\mathbf{Y}\ ^2$	$QM(W_T)$	$QM(W_T)/\hat{\xi}_3$
Resíduo W_E	por subtração	$\ P_{W_E}\mathbf{Y}\ ^2$	$QM(W_E) = \hat{\xi}_3$	
Total V	N	$\ \mathbf{Y}\ ^2$		

Fonte: Adaptado Bailey (2008).

toda a população, ou seja, há diferença significativa entre as linhas e entre as colunas de toda população.

Exemplo de experimento com efeitos aleatórios

Vejamos um exemplo de experimento, em que os efeitos dos blocos são considerados aleatórios. Esse exemplo foi apresentado de Bailey (2008).

Um experimento foi conduzido para comparar o uso de dois corantes e o não uso de corante para proteção de metal. Dez cordões de metal foram utilizados. Cada cordão foi dividido em três partes e os três tratamentos foram aplicados (sem corante A, corante B e corante C). Nesse caso, os cordões podem ser considerados como blocos. Como os cordões de metal são apenas dez de muitos cordões de uma linha de produção, é mais condizente considerá-los como efeitos aleatórios. Depois que os corantes foram aplicados, os cordões foram deixados no tempo por algum período e suas resistências foram medidas. Os dados desse experimento são mostrados na Tabela 5.4

Tabela 5.4 – Dados do experimento cordões de metal.

Cordões	A	B	C	Total de cordões
1	102,4	108,5	106,8	317,7
2	93,7	92,3	96,7	282,7
3	97,4	93,1	100,6	291,1
4	96,1	106,9	101,9	304,9
5	102,5	92,0	103,3	297,8
6	87,8	95,5	94,9	278,2
7	102,6	108,4	106,5	317,5
8	95,2	94,6	101,2	291,0
9	96,9	103,4	111,4	311,7
10	92,1	98,2	92,9	283,2
Total de tratamentos	966,7	992,9	1016,2	2975,8

Fonte: Bailey (2008).

Dos dados temos que:

$$\begin{aligned} \text{Soma de quadrados total} &= 102,4^2 + 108,5^2 + \dots + 92,9^2 \\ &= 296231,92 \end{aligned}$$

$$\begin{aligned} \text{Soma de quadrados da média} &= \frac{2975,8^2}{30} \\ &= 295179,52 \end{aligned}$$

$$\begin{aligned} \text{Soma de quadrados dos tratamentos} &= \frac{(966,7^2 + 992,9^2 + 1016,2^2)}{10} - 295179,52 \\ &= 295302,17 - 295179,52 \\ &= 122,65 \end{aligned}$$

$$\begin{aligned} \text{Soma de quadrados dos cordões} &= \frac{317,7^2 + 282,7^2 + \dots + 283,2^2}{3} - 295179,52 \\ &= 295827,42 - 295179,52 \\ &= 647,90 \end{aligned}$$

$$\begin{aligned} \text{Soma de quadrados dos resíduos} &= 296231,92 - 295179,52 - 122,65 - 647,90 \\ &= 281,85. \end{aligned}$$

Esses dados resultam na Tabela 5.5, que contém a análise de variância. Como F_c para tratamentos é 3,92, e F_t para 2 e 18 graus de liberdade é 3,55, temos que $F_c > F_t$. Logo, rejeita-se a hipótese nula de que não existe diferença entre as médias dos tratamentos e conclui-se que existe efeito de tratamento.

Tabela 5.5 – Tabela de análise de variância quando os cordões tem efeitos aleatórios.

Fonte de variação	Graus de liberdade	Somas de quadrados	Quadrado médio	Estatística <i>F</i>
Média	1	295179,52	295179,52	
Cordões	9	647,90	72,00	
Tratamentos	2	122,65	62,32	3,92
Resíduo	18	281,85	15,66	
Total	30	2962,32		

Fonte: Bailey (2008).

Os resultados da Tabela 5.5 são exatamente os mesmos se considerarmos os cordões como efeitos fixos, o que muda é a interpretação. A hipótese nula testada é de que não existe variabilidade entre os cordões. Nesse caso, para realizar o teste, compare a razão de variâncias $72,00/15,66 = 4,60$. Como esse valor é maior que um rejeita-se a hipótese nula e conclui-se que existe variabilidade entre os cordões. Uma vez que os cordões foram considerados com efeitos aleatórios, essa conclusão pode ser estendida a toda população de cordões, ou seja, toda linha de produção de cordões. Logo, há diferença significativa entre os cordões de toda população.

6 FUNDAMENTOS DA TEORIA DOS DIAGRAMAS DE HASSE APLICADOS À ANÁLISE DE VARIÂNCIA

Nesse capítulo, objetivou-se evidenciar a teoria dos diagramas de Hasse bem como a matemática utilizada nos delineamentos experimentais. Essa teoria está fortemente alicerçada na teoria dos conjuntos. Desse modo, é possível formalizar, de maneira bastante abrangente, a teoria dos delineamentos experimentais.

Bailey (2008) destaca que a utilização dos diagramas de Hasse é pouco difundida no ensino acadêmico. Mas, uma vez aprendida, os estudantes não mais precisam decorar os graus de liberdade de experimentos mais complexos.

Como de costume, a formalização matemática bem como as definições e teoremas, deste capítulo, estão alicerçados em Bailey (2008).

6.1 Fatores, níveis, classes, equivalência e fatores especiais

Considere um conjunto A e uma função f definida em A com contradomínio arbitrário. Sendo esse contradomínio genérico não é possível, por exemplo, somar ou multiplicar funções. Nesse caso, o que caracteriza a função f é que ela define uma partição¹ no conjunto A , formada pelas pré-imagens de f . As propriedades de f ficam totalmente caracterizadas pela partição que ela determina em A .

Por exemplo, dadas duas funções f e g , tais que, $f : A \rightarrow B$ e $g : A \rightarrow C$, tem-se, então, duas partições, $Par(f)$ e $Par(g)$. A intersecção de qualquer subconjunto da partição f com qualquer subconjunto da partição g é outra partição.

As construções de funções que definem partições permitem a formalização matemática da teoria dos delineamentos experimentais. Considere funções com o domínio ou no conjunto de parcelas Ω ou no conjunto de tratamentos \mathfrak{T} . Essas funções são denominadas fatores e denotadas por letras maiúsculas F e G .

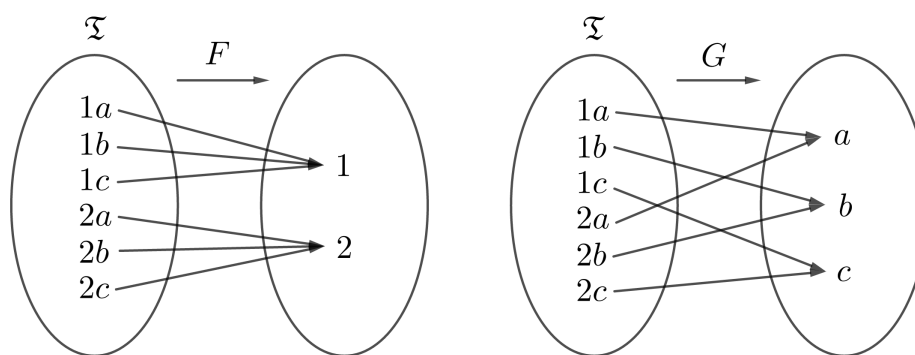
Fatores definidos em Ω são denominados fatores de parcela. Fatores definidos em \mathfrak{T} são denominados fatores de tratamento. Para qualquer fator F no conjunto Ω , temos que $F(\omega)$ é o nível de F que ocorre na parcela ω . Para qualquer fator G no conjunto \mathfrak{T} , temos que $G(i)$ é o nível de G que ocorre no tratamento i .

¹ Veja Definição .0.10 do Apêndice deste trabalho.

Na Seção 3.4, sobre experimentos fatoriais, nos referimos a fatores e seus níveis, sem mencionar funções que definem partições. Agora é possível formalizar, matematicamente, esses conceitos, por meio dessas funções.

Considere o experimento fatorial 2×3 onde os tratamentos são todas as combinações de $\{1, 2\}$ e $\{a, b, c\}$. O conjunto de tratamentos fica, $\mathfrak{T} = \{1a, 1b, 1c, 2a, 2b, 2c\}$. Seja F o fator que associa a cada elemento de \mathfrak{T} um elemento de $\{1, 2\}$ e G o fator que associa a cada elemento de \mathfrak{T} um elemento de $\{a, b, c\}$ (Figura 6.1). Os fatores de tratamentos F e G definem partições no conjunto \mathfrak{T} . Note que F possui dois níveis, 1 e 2, enquanto G possui três níveis, a , b e c .

Figura 6.1 – Fatores de tratamento F e G .



Fonte: Da autora (2019).

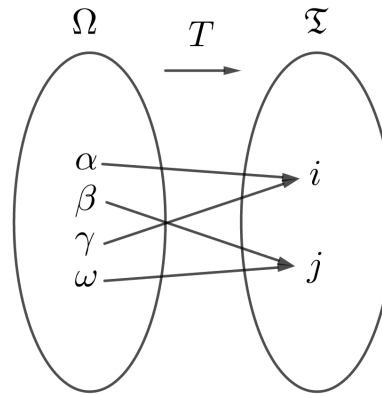
Outro conceito fundamental é o de classe de um fator. Seja F um fator de parcela. A classe de F que contém a parcela β é definida como:

$$F[[\beta]] = \{\omega \in \Omega : F(\omega) = F(\beta)\}.$$

De modo análogo, seja G um fator de tratamento. A classe de G que contém o tratamento j é definida como:

$$G[[j]] = \{i \in \mathfrak{T} : G(i) = G(j)\}.$$

Agora, é possível definir a função de delineamento T , apresentada na Seção 2.2, como um fator de parcela. Para o exemplo exposto naquela seção, cujo diagrama é apresentado novamente na Figura 6.2, as classes são: $T[[\alpha]] = \{\alpha, \gamma\}$ e $T[[\beta]] = \{\beta, \omega\}$. De modo que, $T[[\alpha]]$ é o conjunto de todas as parcelas com o mesmo nível i e $T[[\beta]]$ é o conjunto de todas as parcelas com o mesmo nível j .

Figura 6.2 – Fator de parcela T .

Fonte: Da autora (2019).

Após especificar o que são fatores e suas classes é possível mencionar a definição de equivalência.

Definição 6.1.1. *Sejam F e G fatores no mesmo conjunto. Se toda classe de F for também uma classe de G , então, dizemos que F é equivalente a G , e escrevemos $F \equiv G$.*

Se F e G são fatores de parcela, então $F \equiv G$ se, e somente se, $F[[\omega]] = G[[\omega]]$, para todo $\omega \in \Omega$. Se F e G são fatores de tratamento, então $F \equiv G$ se, e somente se, $F[[i]] = G[[i]]$, para todo $i \in \mathfrak{I}$.

Outra definição importante, com relação a fatores e suas classes, refere-se ao conceito de fator mais "fino" ou fator "precedente".

Definição 6.1.2. *Sejam F e G fatores no mesmo conjunto. Se toda classe de F estiver contida em alguma classe de G , mas $F \neq G$, então, F é mais fino que G , ou G é mais grosso que F , e escreve-se $F \prec G$.*

Se $F \prec G$ ou $F \equiv G$, escreve-se $F \preceq G$.

Se F e G são fatores de parcela, então $F \prec G$ significa que $F[[\omega]] \subset G[[\omega]]$ e $F \preceq G$ significa que $F[[\omega]] \subseteq G[[\omega]]$, para todo $\omega \in \Omega$. De modo análogo, se F e G são fatores de tratamento, então $F \prec G$ significa que $F[[i]] \subset G[[i]]$ e $F \preceq G$ significa que $F[[i]] \subseteq G[[i]]$, para todo $i \in \mathfrak{I}$.

Existem fatores especiais que devem ser mencionados, os quais são apresentados a seguir.

Fator Igualdade: denominado E , é definido por $E(\omega) = \omega$, para todo $\omega \in \Omega$.

Fator Universal: denominado U , é definido por $U(\omega) = 1$, para todo $\omega \in \Omega$.

Fator Uniforme: um fator é dito uniforme se todas as suas classes tiverem o mesmo tamanho.

Note que o fator E possui tantas classes quanto parcelas, já o fator U possui uma única classe que é o próprio Ω .

Observe que para qualquer fator F vale que, $E \preceq F \preceq U$.

Seja F um fator qualquer e n_F o número de níveis de F . Seja F um fator uniforme e k_F o tamanho das classes de F .

Para os fatores E e U , definidos em Ω , temos que $n_E = |\Omega|$ e $n_U = 1$. Os fatores E e U são sempre uniformes, isso porque, $k_E = 1$ e $k_U = |\Omega|$.

Se F é um fator em \mathfrak{T} , então $n_F \times k_F = t$. Se F é um fator uniforme em Ω , então $n_F \times k_F = N$.

No exemplo da Figura 6.1, para o fator de tratamento F , temos que $n_F \times k_F = 2 \times 3 = 6 = t$. No exemplo da Figura 6.2, para o fator de parcela T , temos que $n_T \times k_T = 2 \times 2 = 4 = N$.

Exemplo 6.1

Um exemplo pode deixar os conceitos apresentados anteriormente mais claros. O exemplo apresentado, a seguir, foi adaptado de Bailey (2008).

Um experimento foi conduzido para comparar duas cultivares (Cropper e Melle) combinadas com três quantidades de fertilizantes (0 kg/ha, 80 kg/ha e 160 kg/ha). A área experimental consistia de dois campos, cada um dividido em 2 faixas, sendo que cada faixa consistia de três parcelas. As cultivares foram semeadas em faixas inteiras, porque não é praticável semear em faixas pequenas a não ser à mão. As quantidades diferentes de fertilizantes foram aplicados em cada parcela, porque podem ser aplicados a áreas menores. Veja o croqui desse experimento na Figura 6.3.

Figura 6.3 – Croqui do experimento.

Campo 1				Campo 2			
Cropper 80 ₁	Cropper 160 ₂	Cropper 0 ₃	Faixa 1	Melle 160 ₇	Melle 80 ₈	Melle 0 ₉	Faixa 3
Melle 0 ₄	Melle 160 ₅	Melle 80 ₆	Faixa 2	Cropper 0 ₁₀	Cropper 160 ₁₁	Cropper 80 ₁₂	Faixa 4

Fonte: Adaptado Bailey (2008).

O experimento é constituído por doze parcelas e seis tratamentos. Existem dois campos, cada campo com duas faixas, logo, *campo* e *faixa* são fatores de parcela. Esse exemplo trata-se de um experimento fatorial, uma vez que possui diferentes cultivares combinados com diferentes doses de fertilizante. Logo, *cultivares* e *fertilizantes* são fatores de tratamento.

O fator *campo* possui dois níveis, $n_{campo} = 2$. O fator *faixa* possui quatro níveis, $n_{faixa} = 4$. *Campo* é fator uniforme, porque todas as suas classes possuem seis elementos, $k_{campo} = 6$. *Faixa* também é fator uniforme, porque todas as suas classes possuem três elementos, $k_{faixa} = 3$. Portanto, $n_{campo} \times k_{campo} = n_{faixa} \times k_{faixa} = 12 = N$.

Da mesma forma, os níveis dos fatores *cultivar* e *fertilizantes* são, $n_{cultivar} = 2$ e $n_{fertilizante} = 3$, respectivamente. Ambos são fatores uniformes, de modo que, $k_{cultivar} = 3$ e $k_{fertilizante} = 2$. Portanto, $n_{cultivar} \times k_{cultivar} = n_{fertilizante} \times k_{fertilizante} = 6 = t$.

Por fim, temos que $faixa \prec campo$.

6.2 Ínfimo e supremo

As definições de ínfimo e supremo estão fortemente relacionadas com a teoria dos delineamentos experimentais.

Definição 6.2.1. *O ínfimo de dois fatores F e G no mesmo conjunto é o fator $F \wedge G$ cujas classes são as intersecções não vazias das classes de F com todas as classes de G . O fator $F \wedge G$ satisfaz:*

- i) $F \wedge G \preceq F$ e $F \wedge G \preceq G$;
- ii) Se H é um fator tal que $H \preceq F$ e $H \preceq G$, então $H \preceq F \wedge G$.

No Exemplo 6.1, considere os fatores de parcela *cultivar* e *campo*. A intersecção de cada classe de *campo* com as classes de *cultivar* resulta no fator *faixa*. Portanto, $campo \wedge cultivar = faixa$.

Considere um experimento fatorial com dois fatores arbitrários F e G definidos em \mathfrak{T} . Nesse caso, $F \wedge G = E$, que são todas as combinações dos níveis dos fatores F e G .

Considere, agora, um delineamento em blocos completos, com os fatores *bloco* e *tratamento* definidos em Ω . Nesse caso, $bloco \wedge tratamento = E$, porque cada tratamento ocorre uma vez em cada bloco o que resulta em todas as parcelas.

Definição 6.2.2. *O supremo de dois fatores F e G no mesmo conjunto é o fator $F \vee G$ que satisfaz:*

i) $F \preceq F \vee G$ e $G \preceq F \vee G$;

ii) Se H é um fator tal que $F \preceq H$ e $G \preceq H$, então $F \vee G \preceq H$.

O supremo de dois fatores F e G está associado à partição que contém partições de F e G . Como menciona Bailey (2008), descrever a classe de $F \vee G$ que contém ω não é tão simples como escrever a classe $F \wedge G$.

No Exemplo 6.1, considere o fator de parcela *tratamento*. Vamos escrever *faixa* \vee *tratamento*. Primeiro enumere as parcelas de 1 a 12. Comece na parcela 1, as parcelas na mesma faixa são 1, 2 e 3. As parcelas com o mesmo tratamento que a 1 são as parcelas 1 e 12. As parcelas com o mesmo tratamento que a 2 são as parcelas 2 e 11. As parcelas com o mesmo tratamento que a 3 são as parcelas 3 e 10. As parcelas escritas até o momento foram $\{1, 2, 3, 10, 11, 12\}$. Note que esse conjunto corresponde a duas faixas inteiras (faixa 1 e faixa 4). As parcelas dessas faixas possuem a cultivar Cropper. Portanto, $\text{faixa} \vee \text{tratamento} = \text{cultivar}$.

Para o experimento fatorial com fatores F e G definidos em \mathfrak{X} temos que $F \vee G = U$, porque existe apenas uma combinação de um nível de F com um nível de G .

Para o delineamento em blocos completos, sejam os fatores *bloco* e *tratamento* definidos em Ω . As parcelas em uma classe de *tratamento* contém uma parcela em cada classe de *bloco*, de modo que, $\text{bloco} \vee \text{tratamento}$ pode ter apenas uma classe. Portanto, $\text{bloco} \vee \text{tratamento} = U$.

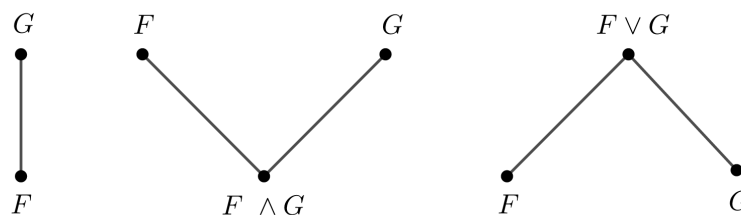
6.3 Diagramas de Hasse e graus de liberdade

Segundo Bailey (2008) as relações existentes entre fatores podem ser representadas por diagramas de Hasse. No desenho do diagrama existe um ponto para cada fator.

Sejam os fatores F e G definidos no mesmo conjunto. Se $F \prec G$, desenhe o ponto de F abaixo do ponto de G e junte-os por uma linha. Nesse caso, o fator F possui mais classes que o fator G .

Se $F \not\prec G$, verifique se existe um ponto $F \wedge G$ ou $F \vee G$. Caso exista $F \wedge G$, desenhe seu ponto abaixo de F e G e junte-os por duas linhas. Caso exista $F \vee G$, desenhe seu ponto acima de F e G e junte-os por duas linhas.

O fator E está sempre na parte inferior, pois, possui tantas classes quanto elementos. Já o fator U está sempre na parte superior, pois, possui uma única classe.

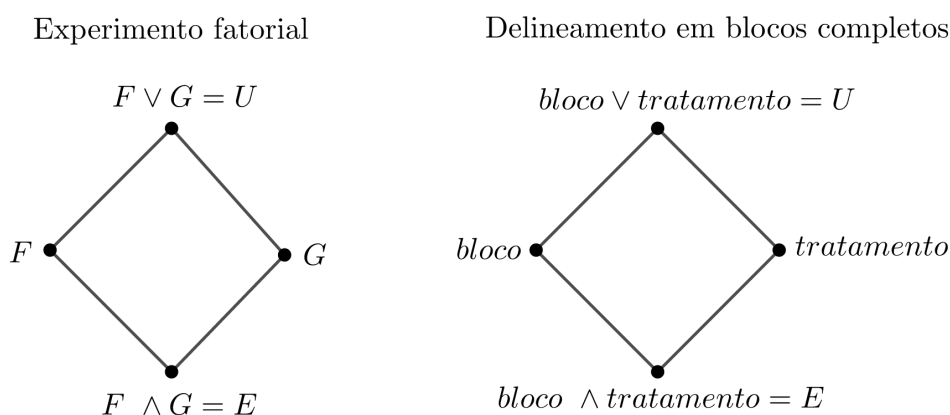
Figura 6.4 – Diagramas de Hasse para $F \prec G$ e $F \not\prec G$.

Fonte: Bailey (2008).

No experimento fatorial, se F e G são fatores de tratamento e nenhum é mais fino que o outro, então, incluímos o ponto $F \wedge G$ no diagrama, a não ser que a interação entre os fatores seja nula. Nesse experimento, é possível incluir também o ponto $F \vee G$. Veja figura à esquerda em 6.5.

No delineamento em blocos completos, para os fatores de parcela *bloco* e *tratamento*, incluímos os pontos $\text{bloco} \wedge \text{tratamento}$ e $\text{bloco} \vee \text{tratamento}$. Veja figura à direita em 6.5.

Figura 6.5 – Diagramas de Hasse para fatores de tratamento e parcela.

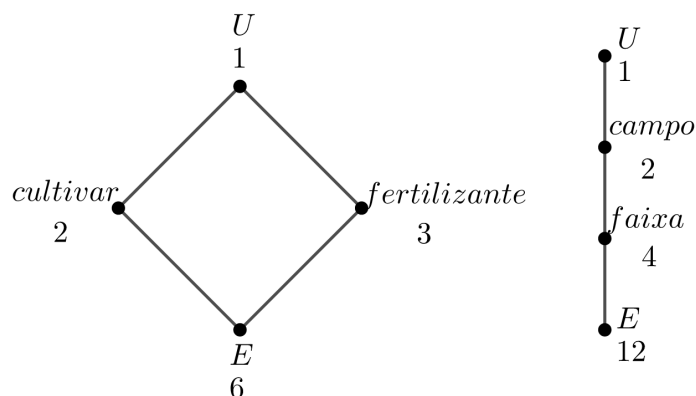


Fonte: Da autora (2019).

Se houver dois fatores, além de U e E , e um for mais fino que o outro, temos o diagrama em série. Agora, se houver dois fatores, além de U e E , e um não for mais fino que o outro temos o diagrama de diamante. É comum nos diagramas de Hasse escrever abaixo de cada fator a quantidade de níveis.

No Exemplo 6.1, para os fatores de tratamento *cultivar* e *fertilizante* o diagrama de Hasse é dado por uma figura de diamante apresentada na Figura 6.6. Para os fatores de parcela *faixa* e *campo* o diagrama de Hasse é dado pela figura em série em 6.6.

Figura 6.6 – Digramas de Hasse para o Exemplo 6.1.



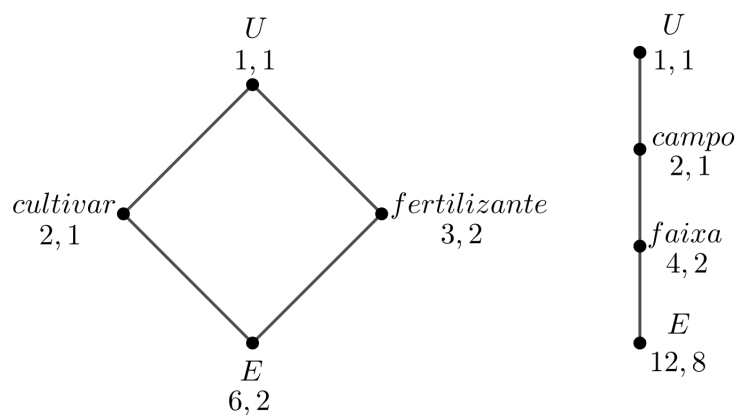
Fonte: Adaptado Bailey (2008).

Existe uma infinidade de diagramas de Hasse que envolvem delineamentos experimentais, esses vão de estruturas simples a estruturas complexas, dependendo do delineamento experimental.

O diagrama de Hasse pode ser utilizado para calcular os graus de liberdade. No diagrama inicie sempre na parte superior pelo fator U e encerre na parte inferior pelo fator E . Para cada fator existente no diagrama, calcule seus graus de liberdade, a partir dos níveis de cada fator. Para tanto, para cada fator do diagrama, subtraia de cada nível os graus de liberdade de todos os fatores acima.

Para o Exemplo 6.1, no diagrama de Hasse os graus de liberdade são colocados à direita dos níveis na Figura 6.7.

Figura 6.7 – Graus de liberdade para o Exemplo 6.1.



Fonte: Adaptado Bailey (2008).

6.4 Subespaços definidos por fatores, ortogonalidade e projeções

O subespaço definido pelo fator F é dado pelo conjunto de vetores que são constantes em cada nível de F . Graficamente, se F possui n_F níveis, as coordenadas do subespaço V_F podem ser descritas como na Figura 6.8.

Figura 6.8 – Coordenadas de V_F .

n_1	n_2	\dots	n_F
<i>iguais</i>	<i>iguais</i>		<i>iguais</i>

Fonte: Da autora (2019).

Seja o conjunto Ω de todas as parcelas de um experimento, lembre-se que a esse conjunto está associado um espaço vetorial N -dimensional $\mathbb{R}^N = V$. Se F é fator de parcela em Ω , as coordenadas de vetores em V são particionadas de modo que partes correspondentes ao mesmo nível sejam iguais.

Particularmente, se U e E são fatores de parcela, o subespaço definido por U é o conjunto de vetores que possui todas as coordenadas constantes V_U , que é o mesmo que V_0 . Já o subespaço definido por E é o espaço todo, que é o mesmo que V .

Para qualquer fator de parcela definido em Ω , se V_F é o subespaço definido pelo fator F é possível se obter a base que gere esse subespaço. Assim, dado um vetor de dados \mathbf{Y} , é possível se obter a projeção de \mathbf{Y} em V_F , que é dada por $P_{V_F}\mathbf{Y}$. O vetor de projeção, em qualquer subespaço definido por um fator, é um vetor de médias, de modo que, parcelas que estão no mesmo nível devem ter as mesmas médias. De modo particular, para $V_0 = V_U$, temos que $P_{V_0}\mathbf{Y}$ é o vetor de média geral.

Como as projeções são vetores de médias, os subespaços definidos por fatores são modelos de médias. Esses modelos são testados por hipóteses na análise de variância, de modo que, todos os subespaços devem ser ortogonais ao subespaço de resíduo.

Com relação a subespaços definidos por fatores, se um fator é mais fino que outro e se o subespaço é definido pelo supremo de dois fatores temos os teoremas que seguem.

Teorema 6.4.1. *Se F e G são fatores no mesmo conjunto e $F \preceq G$, então $V_G \subseteq V_F$.*

Teorema 6.4.2. *Se F e G são fatores no mesmo conjunto, então $V_F \cap V_G = V_{F \vee G}$.*

As demonstrações de ambos os teoremas encontram-se em Bailey (2008).

O subespaço definido por $V_{F \vee G}$ está associado à partição que contenha as partições de F e G . Um vetor em $V_{F \vee G}$ possui coordenadas iguais no mesmo nível de $F \vee G$.

Para fixar ideias, considere os fatores F e G e suas pré-imagens. Um vetor em $V_{F \vee G}^\perp$ é um vetor cuja soma das coordenadas relativas às pré-imagens de F e G é zero (porque $V_{F \vee G}^\perp$ deve ser ortogonal a V_F e V_G). Agora um vetor em $V_F \cap V_{F \vee G}^\perp$, é um vetor cuja soma de suas coordenadas relativas às pré-imagens de F é zero e também a soma das coordenadas relativas a F apenas é zero. O mesmo ocorre para $V_G \cap V_{F \vee G}^\perp$.

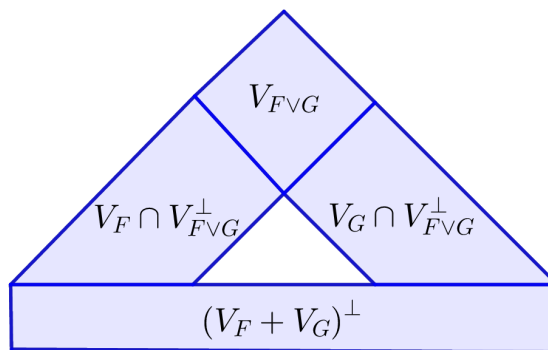
Referente aos subespaços citados anteriormente temos a seguinte definição.

Definição 6.4.1. *Os fatores F e G no mesmo conjunto são ortogonais se, e somente se, o subespaço $V_F \cap V_{F \vee G}^\perp$ é ortogonal ao subespaço $V_G \cap V_{F \vee G}^\perp$.*

Na Figura 6.9, mostram-se os subespaços $V_F \cap V_{F \vee G}^\perp$ e $V_G \cap V_{F \vee G}^\perp$ ortogonais. Nessa figura, é apresentado também o subespaço gerado pelo supremo $V_{F \vee G}$ e subespaço residual $(V_F + V_G)^\perp$.

A ortogonalidade entre $V_F \cap V_{F \vee G}^\perp$ e $V_G \cap V_{F \vee G}^\perp$ garante que vetores nesses subespaços sejam independentes.

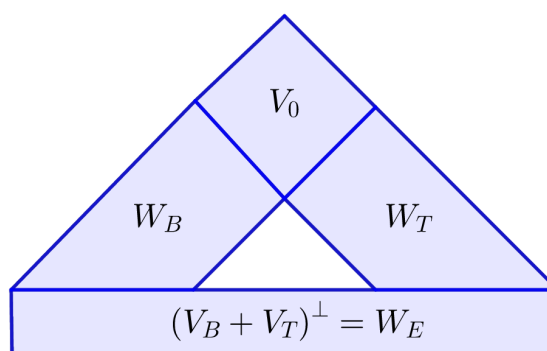
Figura 6.9 – Subespaços ortogonais para os fatores F e G .



Fonte: Da autora (2019).

No delineamento em blocos, sejam os fatores de parcela *bloco* e *tratamento*. Os subespaços definidos por esses fatores são respectivamente V_B e V_T , que são os mesmos definidos na Seção 3.2. Nesse delineamento vimos que $V_B \cap V_T = V_0 = V_U = V_{B \vee T}$. Também na Seção 3.2 vimos que $V_B \cap V_0^\perp = W_B$ e $V_T \cap V_0^\perp = W_T$. Logo, os fatores *bloco* e *tratamento* são ortogonais se, e somente se, W_B é ortogonal a W_T . Na Figura 6.10, mostram-se os subespaços para o delineamento em blocos.

Figura 6.10 – Subespaços ortogonais para o delineamento em blocos.



Fonte: Da autora (2019).

Bailey (2008) destaca que a Definição 6.4.1 não é prática para identificar se dois fatores são ortogonais entre si. Para tanto, as duas condições do teorema a seguir devem ser satisfeitas.

Teorema 6.4.3. *Os fatores F e G no mesmo conjunto são ortogonais se, e somente se:*

$$i) P_{V_F} P_{V_G} \mathbf{Y} = P_{V_G} P_{V_F} \mathbf{Y};$$

$$ii) P_{V_F} P_{V_G} \mathbf{Y} = P_{V_{F \vee G}} \mathbf{Y}.$$

Novamente, a demonstração desse teorema encontra-se em Bailey (2008).

Definição 6.4.2. *Uma coleção de fatores F_1, F_2, \dots, F_n no mesmo conjunto é uma cadeia de fatores se $F_1 \prec F_2 \prec \dots \prec F_n$.*

Os fatores em um diagrama de Hasse de figuras em série formam uma cadeia de fatores. Uma consequência do Teorema 6.4.3 é que uma cadeia de fatores é mutuamente ortogonal.

Exemplo

Vamos usar a Definição 6.4.1, para verificar a ortogonalidade no delineamento em blocos casualizados. Considere o espaço vetorial \mathbb{R}^4 com os fatores *tratamento* e *bloco*. Como já sabemos os subespaços definidos por esses fatores são V_T e V_B . Para um particular experimento, temos que vetores em V_T e V_B são da forma:

$$V_T = \{(x_1, x_2, x_2, x_1) : x_1, x_2 \in \mathbb{R}\}$$

$$V_B = \{(y_1, y_1, y_2, y_2) : y_1, y_2 \in \mathbb{R}\},$$

de modo que $\dim(V_T) = 2$, $\dim(V_B) = 2$. Vetores em $V_T \cap V_B$ são da forma:

$$V_T \cap V_B = \{(\alpha, \alpha, \alpha, \alpha) : \alpha \in \mathbb{R}\} = V_0,$$

de modo que $\dim(V_0) = 1$.

O complemento ortogonal de $V_T \cap V_B$ é caracterizado pelos vetores em que a soma de suas coordenadas resulta em zero. Assim, temos que vetores em $(V_T \cap V_B)^\perp$ são da forma:

$$(V_T \cap V_B)^\perp = \{(a, b, c, d) : a + b + c + d = 0\} = V_0^\perp,$$

de onde vem que:

$$\begin{aligned} V_T \cap (V_T \cap V_B)^\perp &= \{(x_1, x_2, x_2, x_1) : 2x_1 + 2x_2 = 0\} \\ &= \{(x_1, -x_1, -x_1, x_1)\} \\ &= W_T, \end{aligned}$$

$$\begin{aligned} V_B \cap (V_T \cap V_B)^\perp &= \{(y_1, y_1, y_2, y_2) : 2y_1 + 2y_2 = 0\} \\ &= \{(y_1, y_1, -y_1, -y_1)\} \\ &= W_B. \end{aligned}$$

Logo, $W_T \cap W_B = (0, 0, 0, 0)$ e W_T é ortogonal a W_B . Portanto, os fatores *tratamento* e *bloco* são ortogonais.

Exemplo

Vamos usar o Teorema 6.4.3, para verificar a ortogonalidade. Considere o espaço vetorial \mathbb{R}^6 com os fatores F e G . Para um particular experimento os subespaços definidos por esses fatores são V_F e V_G e são da forma:

$$\begin{aligned} V_F &= \{(x_1, x_1, x_2, x_2, x_3, x_3) : x_1, x_2, x_3 \in \mathbb{R}\}, \\ V_B &= \{(y_1, y_2, y_1, y_2, y_3, y_3) : y_1, y_2, y_3 \in \mathbb{R}\}, \end{aligned}$$

de modo que $\dim(V_F) = 3$ e $\dim(V_G) = 3$. Vetores em $V_F \cap V_G$ são da forma:

$$V_F \cap V_G = \{(\alpha, \alpha, \alpha, \alpha, \beta, \beta) : \alpha, \beta \in \mathbb{R}\},$$

de modo que $\dim(V_F \cap V_G) = 2$. As bases para V_F , V_G e $V_F \cap V_G$ são, respectivamente:

$$\{\mathbf{f}_1, \mathbf{f}_2, \mathbf{f}_3\} = \{(1, 1, 0, 0, 0, 0), (0, 0, 1, 1, 0, 0), (0, 0, 0, 0, 1, 1)\},$$

$$\{\mathbf{g}_1, \mathbf{g}_2, \mathbf{g}_3\} = \{(1, 0, 1, 0, 0, 0), (0, 1, 0, 1, 0, 0), (0, 0, 0, 0, 1, 1)\},$$

$$\{\mathbf{v}_1, \mathbf{v}_2\} = \{(1, 1, 1, 1, 0, 0), (0, 0, 0, 0, 1, 1)\}.$$

Seja $\mathbf{Y}^\top = (a, b, c, d, e, f)$ um vetor qualquer em \mathbb{R}^6 . Pelo Teorema 2.2.3 as projeções de \mathbf{Y} nos subespaços V_F , V_G e $V_F \cap V_G$ são:

$$\begin{aligned} P_{V_F} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{f}_1}{\mathbf{f}_1 \cdot \mathbf{f}_1} \right) \mathbf{f}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{f}_2}{\mathbf{f}_2 \cdot \mathbf{f}_2} \right) \mathbf{f}_2 + \left(\frac{\mathbf{Y} \cdot \mathbf{f}_3}{\mathbf{f}_3 \cdot \mathbf{f}_3} \right) \mathbf{f}_3 \\ &= \left(\frac{a+b}{2} \right) \mathbf{f}_1 + \left(\frac{c+d}{2} \right) \mathbf{f}_2 + \left(\frac{e+f}{2} \right) \mathbf{f}_3 + \\ &= \left(\frac{a+b}{2}, \frac{a+b}{2}, \frac{c+d}{2}, \frac{c+d}{2}, \frac{e+f}{2}, \frac{e+f}{2} \right) \\ &= \mathbf{f}, \end{aligned}$$

$$\begin{aligned} P_{V_G} \mathbf{Y} &= \left(\frac{\mathbf{Y} \cdot \mathbf{g}_1}{\mathbf{g}_1 \cdot \mathbf{g}_1} \right) \mathbf{g}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{g}_2}{\mathbf{g}_2 \cdot \mathbf{g}_2} \right) \mathbf{g}_2 + \left(\frac{\mathbf{Y} \cdot \mathbf{g}_3}{\mathbf{g}_3 \cdot \mathbf{g}_3} \right) \mathbf{g}_3 \\ &= \left(\frac{a+c}{2} \right) \mathbf{g}_1 + \left(\frac{b+d}{2} \right) \mathbf{g}_2 + \left(\frac{e+f}{2} \right) \mathbf{g}_3 \\ &= \left(\frac{a+c}{2}, \frac{b+d}{2}, \frac{a+c}{2}, \frac{b+d}{2}, \frac{e+f}{2}, \frac{e+f}{2} \right) \\ &= \mathbf{g}, \end{aligned}$$

$$\begin{aligned} P_{V_F \cap V_G} \mathbf{Y} &= P_{V_{F \cap V_G}} \mathbf{Y} \\ &= \left(\frac{\mathbf{Y} \cdot \mathbf{v}_1}{\mathbf{v}_1 \cdot \mathbf{v}_1} \right) \mathbf{v}_1 + \left(\frac{\mathbf{Y} \cdot \mathbf{v}_2}{\mathbf{v}_2 \cdot \mathbf{v}_2} \right) \mathbf{v}_2 \\ &= \left(\frac{a+b+c+d}{4} \right) \mathbf{v}_1 + \left(\frac{e+f}{2} \right) \mathbf{v}_2 \\ &= \left(\frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{e+f}{2}, \frac{e+f}{2} \right). \end{aligned}$$

Primeiro, vamos verificar o item i) do Teorema 6.4.3, observe que:

$$\begin{aligned}
 P_{V_F}(P_{V_G}\mathbf{Y}) &= P_{V_F}\mathbf{g} \\
 &= \left(\frac{\mathbf{g} \cdot \mathbf{f}_1}{\mathbf{f}_1 \cdot \mathbf{f}_1}\right) \mathbf{f}_1 + \left(\frac{\mathbf{g} \cdot \mathbf{f}_2}{\mathbf{f}_2 \cdot \mathbf{f}_2}\right) \mathbf{f}_2 + \left(\frac{\mathbf{g} \cdot \mathbf{f}_3}{\mathbf{f}_3 \cdot \mathbf{f}_3}\right) \mathbf{f}_3 \\
 &= \left(\frac{a+b+c+d}{4}\right) \mathbf{f}_1 + \left(\frac{a+b+c+d}{4}\right) \mathbf{f}_2 + \left(\frac{e+f}{2}\right) \mathbf{f}_3 \\
 &= \left(\frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{e+f}{2}, \frac{e+f}{2}\right),
 \end{aligned}$$

$$\begin{aligned}
 P_{V_G}(P_{V_F}\mathbf{Y}) &= P_{V_G}\mathbf{f} \\
 &= \left(\frac{\mathbf{f} \cdot \mathbf{g}_1}{\mathbf{g}_1 \cdot \mathbf{g}_1}\right) \mathbf{g}_1 + \left(\frac{\mathbf{f} \cdot \mathbf{g}_2}{\mathbf{g}_2 \cdot \mathbf{g}_2}\right) \mathbf{g}_2 + \left(\frac{\mathbf{f} \cdot \mathbf{g}_3}{\mathbf{g}_3 \cdot \mathbf{g}_3}\right) \mathbf{g}_3 \\
 &= \left(\frac{a+b+c+d}{4}\right) \mathbf{g}_1 + \left(\frac{a+b+c+d}{4}\right) \mathbf{g}_2 + \left(\frac{e+f}{2}\right) \mathbf{g}_3 \\
 &= \left(\frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{a+b+c+d}{4}, \frac{e+f}{2}, \frac{e+f}{2}\right).
 \end{aligned}$$

Logo, $P_{V_F}(P_{V_G}\mathbf{Y}) = P_{V_G}(P_{V_F}\mathbf{Y})$.

Agora, para verificar o item ii) do Teorema 6.4.3, observe que $P_{V_F}(P_{V_G}\mathbf{Y}) = P_{V_F \cap V_G}\mathbf{Y}$.

Portanto, os fatores F e G são ortogonais.

6.5 Delineamentos ortogonais

Como visto, existem fatores de tratamento e fatores de parcela. Fatores definidos em \mathfrak{T} que são ortogonais são chamados de estrutura ortogonal de tratamento. Fatores definidos em Ω que são ortogonais são chamados de estrutura ortogonal de parcela. Essas estruturas, juntamente com a função de delineamento T , são essenciais na constituição dos delineamentos experimentais

Seguem, então, as definições para estruturas ortogonais de tratamento e de parcela.

Estrutura ortogonal de tratamento

Definição 6.5.1. *Um conjunto de fatores \mathcal{G} , não equivalentes no conjunto de tratamentos \mathfrak{T} , é uma estrutura ortogonal de tratamento se:*

- i) $U \in \mathcal{G}$;

- ii) Se $F \in \mathcal{G}$ e $G \in \mathcal{G}$, então $F \vee G \in \mathcal{G}$;
- iii) Se $F \in \mathcal{G}$ e $G \in \mathcal{G}$, então F é ortogonal a G .

Estrutura ortogonal de parcela

Bailey (2008) afirma que estruturas ortogonais de parcela devem satisfazer mais três condições além das apresentadas na estrutura ortogonal de tratamento. Primeiro, todos os fatores devem ser uniformes. Em segundo lugar, E deve ser incluído para obter uma decomposição do espaço todo. Terceiro, se F e G são fatores de parcela então $F \wedge G$ também deve ser.

Definição 6.5.2. Um conjunto de fatores \mathcal{F} , não equivalentes no conjunto de parcelas Ω , é uma estrutura ortogonal de parcela se:

- i) Todo fator F em \mathcal{F} é uniforme;
- ii) $U \in \mathcal{F}$;
- iii) $E \in \mathcal{F}$;
- iv) Se $F \in \mathcal{F}$ e $G \in \mathcal{F}$, então $F \vee G \in \mathcal{F}$;
- v) Se $F \in \mathcal{F}$ e $G \in \mathcal{F}$, então $F \wedge G \in \mathcal{F}$;
- vi) Se $F \in \mathcal{F}$ e $G \in \mathcal{F}$, então F é ortogonal a G .

Note que o fator U deve ser considerado na estrutura ortogonal de tratamento e parcela.

A partir de agora, considere \mathcal{G} como estrutura ortogonal de tratamento e \mathcal{F} como estrutura ortogonal de parcela.

Função de delineamento T

A função de delineamento T já foi mencionada anteriormente, Figura 6.2. Segundo Bailey (2008), uma propriedade da função T é que fatores de tratamento que são ortogonais entre si em \mathfrak{T} devem ser ortogonais entre si em Ω . Iguais repetições dos tratamentos garante que isso aconteça.

De acordo com a mesma autora, para delineamentos ortogonais, se F é um fator de tratamento, é necessário que F seja ortogonal a todos os fatores de parcela. De modo particular,

T também deve ser ortogonal a todos os fatores de parcela. Em geral, se F é fator de tratamento e G é fator de parcela é necessário $F \vee G$ seja fator de tratamento e então $T \preceq F \preceq F \vee G$.

Agora, podemos trazer uma definição geral sobre delineamentos ortogonais.

Definição 6.5.3. *Um delineamento cuja estrutura de tratamento consiste em um conjunto de fatores \mathcal{G} em \mathcal{T} , cuja estrutura de parcela consiste em um conjunto de fatores \mathcal{F} em Ω e cujos tratamentos são alocados a parcelas de acordo com uma função de delineamento T é um delineamento ortogonal se:*

- i) \mathcal{G} é uma estrutura ortogonal de tratamento;
- ii) \mathcal{F} é uma estrutura ortogonal de parcela;
- iii) A função T é tal que:
 - a) Todos os fatores de tratamento em \mathcal{G} permanecem ortogonais entre si em Ω ;
 - b) Se $F \in \mathcal{F}$ e $G \in \mathcal{G}$ então F é ortogonal a G ;
 - c) Se $F \in \mathcal{F}$ e $G \in \mathcal{G}$ então $F \vee G \in \mathcal{G}$.

Dessa forma, um delineamento ortogonal possui fatores de tratamento e fatores de parcela que são ortogonais entre si. Nesse tipo de delineamento, um fator de tratamento é mais grosso que um fator de parcela. Assim, um subespaço gerado por fator de tratamento está contido em um subespaço gerado por fator de parcela.

Para localizar os subespaços de tratamentos que estão contidos em subespaços de parcelas use o teorema seguinte.

Teorema 6.5.1. *Seja F um fator de tratamento em um delineamento ortogonal. Então, existe um único fator de parcela G mais fino ou equivalente a F , de modo que qualquer outro fator de parcela que seja mais fino ou equivalente a F seja mais fino ou equivalente a G . Portanto, $W_F \subseteq W_G$.*

Veja demonstração desse teorema em Bailey (2008).

Assim, os subespaços para delineamentos ortogonais são localizados combinando o diagrama de Hasse para fatores de tratamento e fatores de parcela. De modo que, para cada fator de tratamento, basta encontrar o fator de parcela mais grosso ou equivalente a ele.

6.6 Análise de variância por meio do diagrama de Hasse

Todos os conceitos vistos até o momento são de extrema relevância para a obtenção da análise de variância. Nesta seção é apresentado o esquema da análise de variância² para os delineamentos experimentais vistos até o momento.

Para se obter a análise de variância, por meio de fatores para um delineamento ortogonal que atende às condições da Definição 6.5.3, primeiro obtenha os diagramas de Hasse e em seguida o esquema da análise de variância.

Para obter os diagramas de Hasse siga os seguintes passos:

- 1) Localize a estrutura ortogonal de tratamento e faça o diagrama de Hasse correspondente. Neste diagrama escreva o nome de cada fator, coloque a quantidade de níveis de cada fator e ao lado dos níveis os graus de liberdade;
- 2) Localize a estrutura ortogonal de parcela e faça o diagrama de Hasse correspondente. Nesse diagrama, escreva o nome de cada fator, coloque a quantidade de níveis de cada fator e ao lado dos níveis os graus de liberdade;
- 3) Combine os diagramas de Hasse de 1) e 2) em um diagrama. Nesse momento, verifique quais fatores de parcela são mais finos que fatores de tratamento. De modo que, o Teorema 6.5.1 ajuda a decidir quais subespaços gerados por fatores de tratamento estão contidos em quais subespaços gerados por fatores de parcela.

Para se obter o esquema de análise de variância, a partir do diagrama de Hasse combinado, siga os seguintes passos:

- 1) Insira uma coluna para os fatores de parcela. Nessa coluna coloque uma linha para cada fator de parcela do experimento. De modo que, a primeira linha corresponde ao fator U e a última corresponde ao fator E .
- 2) Insira uma coluna para fatores de tratamento, que são as fontes de variação. Nessa coluna coloque uma linha para cada fator de tratamento dentro da linha do fator de parcela correspondente. Nesse momento, verifique qual fator de parcela é mais fino o que um fator de tratamento. Note que a linha para o fator de parcela que contém fatores de tratamento

² O esquema da análise de variância possui consiste em: fatores de parcela; fatores de tratamentos, que são as fontes de variação e ; graus de liberdade.

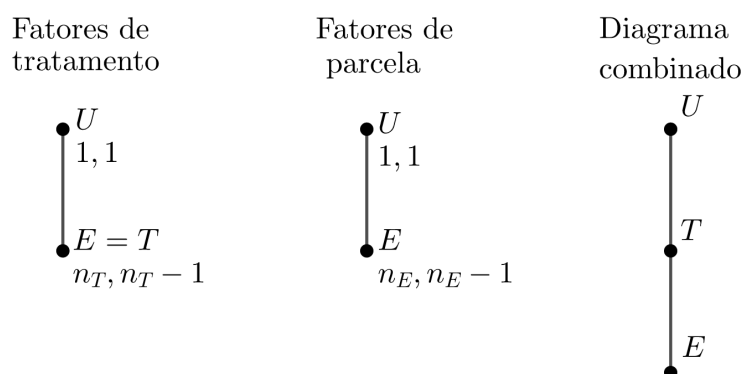
deve incluir, além das linhas para fatores de tratamentos, uma linha para o resíduo e outra para o total.

- 3) Insira uma coluna para os graus de liberdade. Essa coluna é dividida em duas, à direita coloque os graus de liberdade para fatores de parcela e à esquerda coloque os graus de liberdade para fatores de tratamento.

Agora é possível apresentar o esquema de análise de variância para os delineamentos apresentados, neste texto, por meio do diagrama de Hasse combinado.

No delineamento inteiramente casualizado, sejam os fatores de tratamento U e $E = T$, os subespaços gerados por esses fatores são V_0 e V_T . Sejam agora fatores de parcela U e E , os subespaços gerados por esses fatores são V_0 e V . Os diagramas de Hasse, para esses fatores são apresentados na Figura 6.11 e o esquema da análise de variância na Tabela 6.1. Note que as fontes de variação tratamento e resíduo correspondem aos subespaços $V_T \cap V_0^\perp = W_T$ e $V_T^\perp = W_E$ respectivamente, esses são ortogonais pela Definição 6.4.1.

Figura 6.11 – Diagramas de Hasse para o delineamento inteiramente casualizado.



Fonte: Da autora (2019).

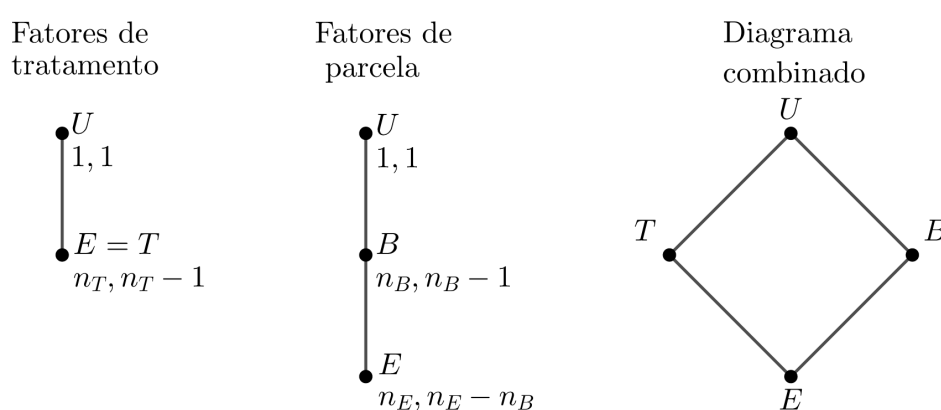
Tabela 6.1 – Esquema de análise de variância para delineamento inteiramente casualizado por meio do diagrama de Hasse.

Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Parcelas	Tratamento	$n_T - 1$
	Resíduo	$n_E - n_T$
	Total	$n_E - 1$
Total		n_E

Fonte: Da autora (2019).

Para o delineamento em blocos casualizados, considere os fatores de tratamento U e $E = T$, os subespaços gerados por esses fatores são V_0 e V_T . Considere também os fatores de parcela U , $\text{bloco} = B$ e E , os subespaços gerados por esses fatores são, respectivamente, V_0 , V_B e V . Os diagramas de Hasse para esses fatores são apresentados na Figura 6.12 e o esquema da análise de variância na Tabela 6.2. Note que as fontes de variação bloco, tratamento e resíduo correspondem aos subespaços $V_B \cap V_0^\perp = W_B$, $V_T \cap V_0^\perp = W_T$ e $(V_T + V_B)^\perp = W_E$ respectivamente, que pela Definição 6.4.1 são ortogonais.

Figura 6.12 – Diagramas de Hasse para o delineamento em blocos casualizados.



Fonte: Da autora (2019).

Tabela 6.2 – Esquema de análise de variância para delineamento em blocos casualizado por meio do diagrama de Hasse.

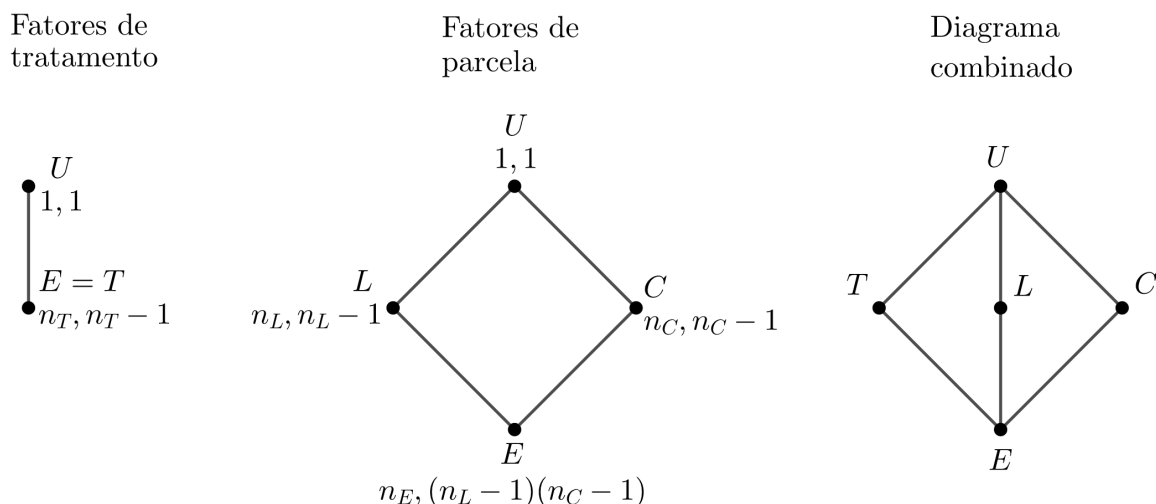
Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Bloco	Bloco	$n_B - 1$
Parcelas	Tratamento	$n_T - 1$
	Resíduo	$(n_B - 1)(n_T - 1)$
	Total	$n_E - n_B$
Total		n_E

Fonte: Da autora (2019).

Para o delineamento em linhas e colunas, admita os fatores de tratamento U e $E = T$, os subespaços gerados por esses fatores são V_0 e V_T . Admita agora os fatores de parcela U , $\text{linha} = L$, $\text{coluna} = C$ e E , os subespaços gerados por esses fatores são, respectivamente V_0 , V_L , V_C e V . Os diagramas de Hasse para esses fatores são apresentados na Figura 6.13 e o esquema da análise de variância na Tabela 6.3. Veja que as fontes de variação linha, coluna,

tratamento e resíduo correspondem aos subespaços $V_L \cap V_0^\perp = W_L$, $V_C \cap V_0^\perp = W_C$, $V_T \cap V_0^\perp = W_T$ e $(V_L + V_C + V_T)^\perp = W_E$ respectivamente, que pela Definição 6.4.1 são ortogonais.

Figura 6.13 – Diagramas de Hasse para o delineamento em linhas e colunas.



Fonte: Da autora (2019).

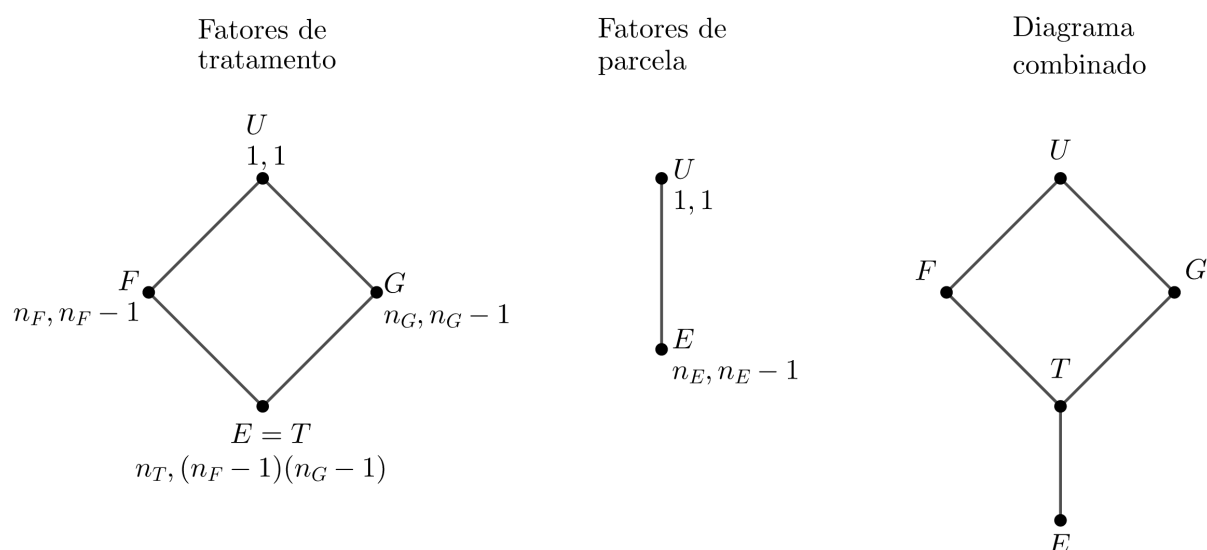
Tabela 6.3 – Esquema de análise de variância para delineamento em linhas e colunas por meio do diagrama de Hasse.

Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Linha	Linha	$n_L - 1$
Coluna	Coluna	$n_C - 1$
Parcelas	Tratamento	$n_T - 1$
	Resíduo	$n_E - n_L - n_C - n_T + 2$
	Total	$(n_L - 1)(n_C - 1)$
Total		n_E

Fonte: Da autora (2019).

Agora para o experimento fatorial, planejado no delineamento inteiramente casualizado, considere os fatores de tratamento U , F , G e $E = T$, os subespaços gerados por esses fatores são V_0 , V_F , V_G e V_T . Considere também os fatores de parcela U e E , os subespaços gerados por esses fatores são V_0 e V . Os diagramas de Hasse para esses fatores são apresentados na Figura 6.14 e o esquema da análise de variância na Tabela 6.4. Veja que as fontes de variação F , G , tratamento e resíduo correspondem aos subespaços $V_F \cap V_0^\perp = W_F$, $V_G \cap V_0^\perp = W_G$, $V_T \cap V_0^\perp = W_T$ e $V_T \cap (V_F + V_G)^\perp = W_{F \wedge G}$ respectivamente, sendo que esses subespaços são ortogonais pela Definição 6.4.1.

Figura 6.14 – Diagramas de Hasse para o experimento fatorial.



Fonte: Da autora (2019).

Tabela 6.4 – Esquema de análise de variância para experimento fatorial por meio do diagrama de Hasse.

Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Parcelas	F	$n_F - 1$
	G	$n_G - 1$
	Tratamento	$(n_F - 1)(n_G - 1)$
	Resíduo	$n_E - n_T$
Total		$n_E - 1$

Fonte: Da autora (2019).

Após obter o esquema de análise de variância, as projeções nos subespaços gerados por fatores são obtidas como de costume. Essas projeções são vetores de médias utilizados para realizar os testes. Sendo os subespaços ortogonais, vetores relativos a esses subespaços são independentes. Assim, é possível realizar o teste entre razões de χ^2 .

6.6.1 Exemplos de experimentos em parcelas subdivididas

A análise de variância, por meio do diagrama de Hasse, pode ser útil em delineamentos mais complexos. Por exemplo, existem situações experimentais em que os fatores de tratamentos estão distribuídos em fatores de parcelas distintos. Nesse caso, a análise de variância possui mais de uma fonte de variação residual.

Na prática, esses experimentos são denominados de experimentos em parcelas subdivididas. As relações definidas por fatores juntamente com o diagrama de Hasse podem ajudar com esses experimentos.

A seguir, são apresentando alguns exemplos de delineamentos ortogonais, em parcelas subdivididas, propostos por Bailey (2008).

Exemplo 6.6.1: continuação do Exemplo 6.1

Para se obter o esquema da análise de variância, combinamos os diagramas de Hasse para fatores de tratamento e fatores de parcela em um diagrama. Para se obter esse diagrama observe que:

- $campo \vee cultivar$ e $campo \vee fertilizante$ são iguais a U ;
- $fertilizante \vee campo = U$, porque em cada campo ocorrem todas as doses de fertilizantes;
- $T \vee faixa = cultivar$, porque em cada faixa ocorre uma cultivar e todas as doses de fertilizantes;
- $faixa \prec cultivar$, porque as faixas possuem todas as cultivares;
- $faixa \prec campo$, porque as faixas ocorrem nos campos;
- $T \prec fertilizante$, porque os tratamentos possuem todas as doses de fertilizantes;
- $E \prec T \prec fertilizante$.

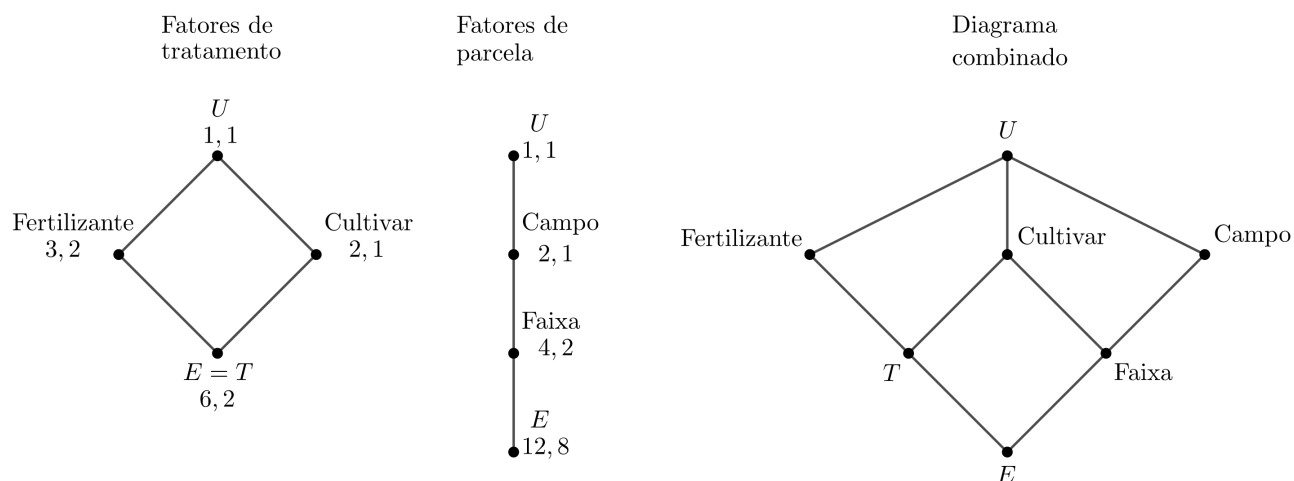
O que resulta no diagrama combinado da Figura 6.15.

Como destaca Bailey (2008) temos que $W_{cultivar} \subseteq W_{faixa}$, $W_{fertilizante} \subseteq W_E$ e também $W_T \subseteq W_E$. Assim, o diagrama de Hasse combinado resulta no esquema de análise de variância da Tabela 6.5.

Exemplo 6.6.2

Um fabricante de eletrodomésticos deseja saber qual a melhor combinação de temperatura de lavagem e temperatura de secagem para lençóis de algodão, de tal maneira que esses não amarrotam no final da lavagem. Ele quer comparar quatro temperaturas de lavagem e três temperaturas de secagem. No experimento, ele usa oito máquinas de lavar e seis secadoras.

Figura 6.15 – Diagramas de Hasse para o Exemplo 6.6.1.



Fonte: Adaptado Bailey (2008).

Tabela 6.5 – Esquema de análise de variância para o Exemplo 6.6.1.

Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Campo	campo	1
Faixa	Cultivar	1
	Resíduo	1
	Total	2
Parcelas	Fertilizantes	2
	<i>T</i>	2
	Resíduo	6
	Total	8
Total		12

Fonte: Da autora (2019).

Primeiro, em cada máquina são alocados seis lençóis, resultando num total de 48 lençóis. Nesse momento, as temperaturas das máquinas de lavar são escolhidas aleatoriamente, de modo que duas lavadoras são executadas com as mesmas temperaturas.

Em seguida, após a lavagem, os seis lençóis de cada máquina são alocados, um em cada secadora, aleatoriamente. Nesse momento, as temperaturas de secagem são escolhidas, aleatoriamente, nas secadoras, de modo que duas secadoras são executadas com as mesmas temperaturas.

Após a lavagem, os 48 lençóis são pontuados por especialistas, pelo quanto estão amarrados..

Considere as quatro temperaturas de lavagem (A, B, C e D) e as três temperaturas de secagem (a, b e c). Note que o experimento trata-se de um fatorial com 12 tratamentos. Veja o croqui desse experimento dado na Figura 6.16.

Figura 6.16 – Croqui do experimento para o exemplo 6.6.2.

Máquina de lavar	1	2	3	4	5	6	7	8	
Temperatura	3	4	2	3	4	1	1	2	
Secadora	Temperatura								
1	1	<i>Ca</i>	<i>Da</i>	<i>Ba</i>	<i>Ca</i>	<i>Da</i>	<i>Aa</i>	<i>Aa</i>	<i>Ba</i>
2	3	<i>Cc</i>	<i>Dc</i>	<i>Bc</i>	<i>Cc</i>	<i>Dc</i>	<i>Ac</i>	<i>Ac</i>	<i>Bc</i>
3	2	<i>Cb</i>	<i>Db</i>	<i>Bb</i>	<i>Cb</i>	<i>Db</i>	<i>Ab</i>	<i>Ab</i>	<i>Bb</i>
4	3	<i>Cc</i>	<i>Dc</i>	<i>Bc</i>	<i>Cc</i>	<i>Dc</i>	<i>Ac</i>	<i>Ac</i>	<i>Bc</i>
5	2	<i>Cb</i>	<i>Db</i>	<i>Bb</i>	<i>Cb</i>	<i>Db</i>	<i>Ab</i>	<i>Ab</i>	<i>Bb</i>
6	1	<i>Ca</i>	<i>Da</i>	<i>Ba</i>	<i>Ca</i>	<i>Da</i>	<i>Aa</i>	<i>Aa</i>	<i>Ba</i>

Fonte: Adaptado Bailey (2008).

Os fatores de tratamento são: U , temperatura de lavagem = TL , temperatura de secagem = TS e $E = T$. O diagrama de Hasse para fatores de tratamentos é o usual para experimentos fatoriais.

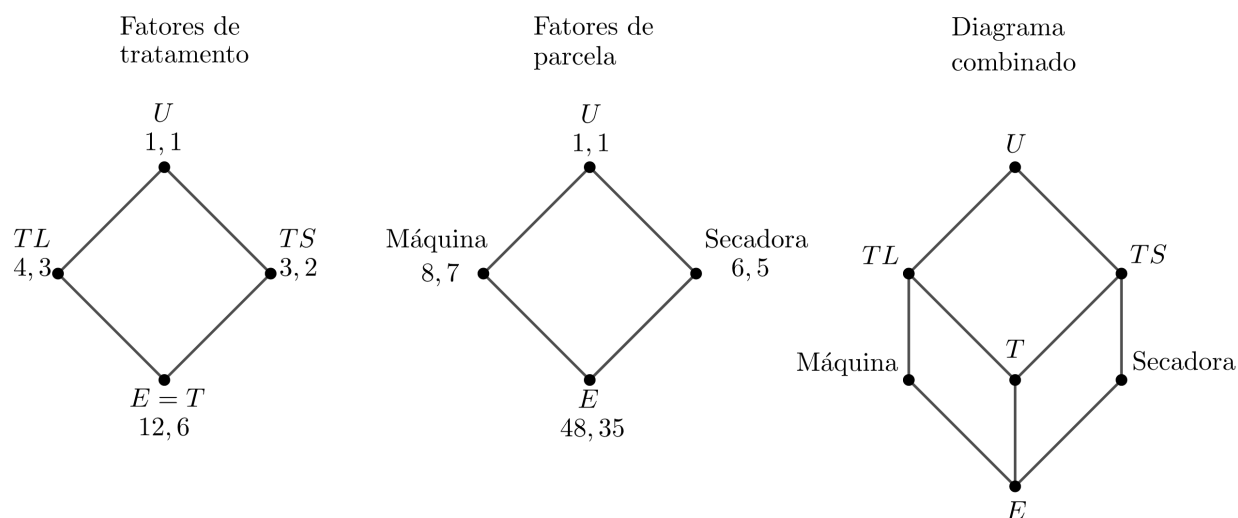
Os fatores de parcela são: U , máquina, secadora e E . Como as parcelas são combinações de máquinas com secadores, o diagrama de Hasse para fatores de parcela é equivalente ao delineamento em linhas e colunas.

Nesse experimento note que:

- $TL \vee TS = U$;
- $máquina \vee T = TL$, porque em cada máquina ocorrem todas as temperaturas de lavagem;
- $secadora \vee T = TS$, porque em cada secadora ocorrem todas as temperaturas de secagem;
- $TL \wedge TS = T$, porque os tratamentos são combinações de temperaturas de lavagem e secagem;
- $E \prec T$.

Os diagramas referentes aos fatores de tratamento e de parcela, bem como o diagrama de Hasse combinado são apresentados, na Figura 6.17.

Figura 6.17 – Diagramas de Hasse para o exemplo 6.6.2.



Fonte: Adaptado Bailey (2008).

Por meio do diagrama de Hasse combinado tem-se o esquema de análise de variância mostrado na Tabela 6.6. Em relação aos subespaços desse experimento, perceba que $W_{TL} \subseteq W_{máquina}$, $W_T \subseteq W_E$ e $W_{secadora} \subseteq W_{TS}$.

Tabela 6.6 – Esquema de análise de variância para o exemplo 6.6.2.

Fatores de parcela	Fontes de variação	Graus de liberdade
Média	Média	1
Máquina	TL	3
	Resíduo	4
	Total	7
Secadora	TS	2
	Resíduo	3
	Total	5
Parcelas	T	6
	Resíduo	26
	Total	35
Total		48

Fonte: Adaptado Bailey (2008).

7 METODOLOGIA

Neste capítulo, é apresentada a metodologia utilizada no trabalho.

De acordo com França e Vasconcelos (2007), uma dissertação deve abordar um tema único, exige investigações próprias na área de especialização e métodos específicos. Nesse sentido, o tema único abordado refere-se aos delineamentos experimentais apresentados nas Seções 3.1, 3.2 e 3.3. As investigações próprias referem-se à abordagem geométrica pouco difundida na literatura, bem como as representações geométricas, sendo algumas inéditas, sobre o tema abordado.

No presente trabalho utilizou-se a metodologia de pesquisa bibliográfica, alicerçada no livro de Bailey (2008). "[...] a pesquisa bibliográfica implica em um conjunto ordenado de procedimentos de busca por soluções, atento ao objeto de estudo, e que, por isso, não pode ser aleatório"(LIMA; MIOTO, 2007, p.38). Para as autoras a pesquisa bibliográfica é um procedimento metodológico que se oferece ao pesquisador como uma possibilidade na busca de soluções para seu problema de pesquisa. De acordo com as mesmas autoras, trabalhar com a pesquisa bibliográfica significa realizar um movimento incansável de compreensão dos objetivos.

Nessa perspectiva, procurou-se focar nos objetivos específico e geral do trabalho, os quais são, respectivamente: abordar os tópicos de Bailey (2008), além de explorar ao máximo aspectos geométricos e disponibilizar um texto amplo com valor didático.

Para atingir o objetivo específico do trabalho, o livro de Bailey (2008) foi estudado minuciosamente. Durante os estudos, foram feitas representações geométricas de conceitos abordados, tais como: espaços e subespaços vetoriais, vetores de projeção, realizações de variáveis aleatórias e testes de hipóteses da análise de variância. Vale destacar que essas representações geométricas foram feitas com o uso do software Geogebra.

Para atingir o objetivo geral do trabalho, as demonstrações dos teoremas não apresentadas por Bailey (2008) foram feitas cautelosamente, bem como algumas passagens algébricas. Nesse momento, outros materiais foram consultados. Vale ressaltar que, durante o texto, são apresentadas definições e teoremas que se encontram no Apêndice deste trabalho, com o objetivo de situar o leitor no texto.

A partir do texto escrito, baseado em Bailey (2008), entende-se que os objetivos do trabalho foram alcançados.

8 RESULTADOS

Os resultados, deste trabalho, referem-se a alcançar o objetivo específico de abordar os tópicos de Bailey (2008), além de explorar ao máximo aspectos geométricos e o objetivo geral de disponibilizar um texto amplo com valor didático.

Os resultados deste trabalho encontram-se ao longo do texto, nos Capítulos 2, 3, 4, 5 e 6. A seguir, para cada objetivo, são descritos em cada capítulo como se entende que os objetivos foram alcançados.

Objetivo específico

No Capítulo 2, as representações geométricas de vetores, espaços e subespaços vetoriais e realizações do vetor aleatório Y permitem explorar aspectos geométricos importantes. Os conceitos estatísticos referentes à análise de variância, do ponto de vista geométrico, permitem uma visão pouco difundida na literatura sobre o tema.

No Capítulo 3, novamente, as representações geométricas permitem o enfoque geométrico. A análise de variância apresentada do ponto de vista geométrico é explorada, bem como as representações de vetores e subespaços para delineamentos experimentais. Neste capítulo, a interpretação geométrica dos testes de hipóteses são inéditas e permitem uma componente visual e também maior compreensão no modo de lidar com as hipóteses envolvidas na análise de variância.

Objetivo geral

No capítulo 2, demonstrações dos teoremas não expostos por Bailey (2008) são apresentadas o que torna o texto completo e didático.

No Capítulo 3, são apresentadas passagens algébricas com relação às somas de quadrados e dimensões dos subespaços da análise de variância e também demonstrações de teoremas não expostos por Bailey (2008). Isso com o intuito de disponibilizar um texto completo e didático.

No capítulo 4, a linguagem geométrica proposta por Bailey (2008) é comparada com a abordagem usual de modelos lineares, o que torna o texto completo.

No capítulo 5, a maneira que a análise de variância na presença de efeitos aleatórios é apresentada com relação aos autoespaços da matriz de covariâncias Σ permite uma compreensão mais didática de como os testes são realizados.

No Capítulo 6, a teoria dos diagramas de Hasse permite uma compreensão formal dos experimentos em parcelas subdivididas. Nesse capítulo a formalização matemática em termos de fatores e a construção correspondente dos diagramas de Hasse simplificam, de forma didática, a teoria dos delineamentos experimentais.

Na maioria dos capítulos, deste trabalho, são apresentados exemplos com o intuito de abordar os assuntos expostos, o que colabora para que um texto se torne didático. Vale destacar que os exemplos não citados são de autoria própria.

Por fim, pode-se dizer que a abordagem geométrica dos delineamentos experimentais, apresentou-se como uma ferramenta didática. Sendo uma maneira interessante de apresentar a teoria dos delineamentos experimentais, análise de variância e testes de hipóteses relacionados.

9 CONCLUSÃO

Ao finalizar este trabalho, é possível concluir que o objetivo específico de abordar os tópicos de Bailey (2008), além de explorar o máximo dos aspectos geométricos, foram alcançados. Isso, porque, em todo o texto, aspectos geométricos foram levantados e também representações geométricas de vetores, espaços e subespaços vetoriais e testes de hipóteses foram evidenciados. Vale destacar que as figuras ilustrativas relacionadas aos testes de hipóteses são inéditas na literatura.

O objetivo geral de disponibilizar um texto amplo com valor didático também foi alcançado. Isso, porque, o texto apresenta demonstrações não expostas por Bailey (2008), passagens algébricas importantes no decorrer do texto e exemplos para enfatizar conceitos apresentados, sendo alguns de autoria própria.

10 REFERÊNCIAS

- ADÃO, A. S. **Introdução a teoria geométrica dos delineamentos experimentais**. 2011. 78 p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) - Universidade Federal de Lavras, Lavras, 2011.
- BAILEY, R. A. **Design of comparative experiments**. New York: Cambridge University Press, 2008. 330 p.
- BARBIN, D. **Componente de variância: teoria e aplicações**. 2. Ed. Piracicaba: Falaq, 1993. 120 p.
- BOLDRINI, J. L. et al. **Álgebra linear**. São Paulo: Harper & Row do Brasil, 1980. 371 p.
- CASELLA, G.; BERGER, R. L. **Inferência estatística**. São Paulo: Cengage Learning, 2010. 612 p.
- FRANÇA, J. L.; VASCONCELOS, A. C. de. **Manual para normalização de publicações técnico-científicas**. 8. ed. rev. e ampl. Belo Horizonte: Ed. UFMG, 2007. 255 p.
- LIMA, T. C. S. D.; MIOTO, R. C. T. **Procedimentos metodológicos na construção do conhecimento científico: a pesquisa bibliográfica**. Revista Katálysis, Florianópolis, v. 10, nesp., p. 37-45, 2007.
- PIMENTEL-GOMES, F. P.; GARCIA, C. H. **Estatística aplicada a experimentos agrônômicos e florestais: exposição com exemplos e orientações para uso de aplicativos**. Piracicaba: FEALQ, 2002. 309 p.
- RENCHER, A. C.; SCHAALJE, G. B. **Linear models in statistics**. New York: J. Wiley & Sons, 2008. 688 p.
- SANTOS, R. J. **Álgebra linear e aplicações**. Belo Horizonte: Imprensa Universitária da UFMG, 2011. 516 p.

APÊNDICE A – DEFINIÇÕES E TEOREMAS RELEVANTES

Definição .0.1. Dizemos que um conjunto $V \neq \emptyset$, munido de duas operações, uma soma e uma multiplicação por escalar:

0) Se $\mathbf{v}, \mathbf{w} \in V$, então $\mathbf{v} + \mathbf{w} \in V$;

0') Se $\mathbf{v} \in V$ e $\alpha \in \mathbb{R}$, então $\alpha\mathbf{v} \in V$;

é um espaço vetorial sobre \mathbb{R} se satisfaz os seguintes axiomas:

1) Para todos os $\mathbf{v}, \mathbf{w} \in V$, $\mathbf{v} + \mathbf{w} = \mathbf{w} + \mathbf{v}$;

2) Para todos os $\mathbf{v}, \mathbf{w}, \mathbf{u} \in V$, $\mathbf{v} + (\mathbf{w} + \mathbf{u}) = (\mathbf{v} + \mathbf{w}) + \mathbf{u}$;

3) Existe um elemento $\mathbf{0} \in V$, tal que $\mathbf{v} + \mathbf{0} = \mathbf{0} + \mathbf{v} = \mathbf{v}$, para todo $\mathbf{v} \in V$;

4) Para cada $\mathbf{v} \in V$, existe um elemento $-\mathbf{v} \in V$ tal que $\mathbf{v} + (-\mathbf{v}) = \mathbf{0}$;

5) Para todo $\mathbf{v} \in V$ e todos escalares α e β , $\alpha(\beta\mathbf{v}) = (\alpha\beta)\mathbf{v}$;

6) Para todos os $\mathbf{v}, \mathbf{w} \in V$ e todo escalar α , $\alpha(\mathbf{v} + \mathbf{w}) = \alpha\mathbf{v} + \alpha\mathbf{w}$;

7) Para todo $\mathbf{v} \in V$ e todos os escalares α e β , $(\alpha + \beta)\mathbf{v} = \alpha\mathbf{v} + \beta\mathbf{v}$;

8) Para todo $\mathbf{v} \in V$, $1\mathbf{v} = \mathbf{v}$.

(SANTOS, 2011, p.5).

Definição .0.2. Dado um espaço vetorial V , um subconjunto W , não vazio, será um subespaço vetorial de V se:

i) Para quaisquer $\mathbf{u}, \mathbf{v} \in W$ tivermos $\mathbf{u} + \mathbf{v} \in W$;

ii) Para quaisquer $\alpha \in \mathbb{R}$, $\mathbf{u} \in W$ tivermos $\alpha\mathbf{u} \in W$.

(BOLDRINI et. al, 1980, p.106).

Definição .0.3. Sejam V um espaço vetorial, $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$ vetores em V e a_1, a_2, \dots, a_N números reais. Então, o vetor

$$\mathbf{v} = a_1\mathbf{v}_1 + a_2\mathbf{v}_2 + \dots + a_N\mathbf{v}_N$$

é uma combinação linear de $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_N$.

(BOLDRINI et. al, 1980, p.112).

Definição .0.4. De todo vetor $\mathbf{v} \neq \mathbf{0}$ de V é possível obter um vetor unitário tomando:

$$\mathbf{u} = \frac{\mathbf{v}}{\|\mathbf{v}\|}.$$

(BOLDRINI et. al, 1980, p.227).

Definição .0.5. Diz-se que uma base $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ de V é base ortogonal se $\mathbf{v}_i \cdot \mathbf{v}_j = 0$ para $i \neq j$, isto é, os vetores da base são dois a dois ortogonais, (Boldrini et. al, 1980, p.225).

Definição .0.6. Seja V um espaço vetorial com produto interno. Diz-se que uma base $\{\mathbf{v}_1, \dots, \mathbf{v}_N\}$ de V é ortonormal se for ortogonal e cada vetor for unitário, isto é

$$\mathbf{v}_i \cdot \mathbf{v}_j = \begin{cases} 0, & \text{se } i \neq j \\ 1, & \text{se } i = j \end{cases}$$

(BOLDRINI et. al, 1980, p.229).

Definição .0.7. Sejam W_1 e W_2 dois subespaços de um espaço vetorial V .

a) Definimos a soma dos subespaços, $W_1 + W_2$, como sendo o conjunto de todos os vetores de V que são soma de um elemento de W_1 com um elemento de W_2 , ou seja,

$$\begin{aligned} W_1 + W_2 &= \{\mathbf{v}_1 + \mathbf{v}_2 \mid \mathbf{v}_1 \in W_1 \text{ e } \mathbf{v}_2 \in W_2\} \\ &= \{\mathbf{v} \in V \mid \mathbf{v} = \mathbf{v}_1 + \mathbf{v}_2 \text{ com } \mathbf{v}_1 \in W_1 \text{ e } \mathbf{v}_2 \in W_2\}. \end{aligned}$$

b) Se o espaço V é tal que

$$\begin{aligned} V &= W_1 + W_2 \text{ e} \\ W_1 \cap W_2 &= \{\mathbf{0}\}, \end{aligned}$$

dizemos que V é soma direta de W_1 e W_2 e denotamos por $V = W_1 \oplus W_2$.

(SANTOS, 2011, p.26).

Teorema .0.1. *Se Y é um vetor aleatório com média $\boldsymbol{\tau}$ e matriz de variâncias e covariâncias $\boldsymbol{\Sigma}$ e A é uma matriz simétrica de constantes então:*

$$E(\mathbf{Y}^T \mathbf{A} \mathbf{Y}) = \text{tr}(\mathbf{A} \boldsymbol{\Sigma}) + \boldsymbol{\tau}^T \mathbf{A} \boldsymbol{\tau}.$$

(RENCHER, 2008, p. 107, tradução nossa).

Definição .0.8. *Sejam V e W espaços vetoriais. Uma transformação linear é uma função de V em W , $F : V \rightarrow W$, que satisfaz as seguintes condições:*

i) *Quaisquer que sejam \mathbf{u} e \mathbf{v} em V ,*

$$F(\mathbf{u} + \mathbf{v}) = F(\mathbf{u}) + F(\mathbf{v});$$

ii) *Quaisquer que sejam $k \in \mathbb{R}$ e $\mathbf{u} \in V$,*

$$Fk(\mathbf{u}) = kF(\mathbf{u}).$$

(BOLDRINI et. al, 1980, p.144).

Definição .0.9. *Seja A uma matriz quadrada e \mathbf{v} um vetor não nulo no domínio de A . Dizemos que \mathbf{v} é um autovetor de associado ao autovalor $\alpha \in \mathbb{R}$, se:*

$$A\mathbf{v} = \alpha\mathbf{v}.$$

Teorema .0.2. *Autovetores, associados a autovalores distintos, de matrizes simétricas são sempre ortogonais.*

Definição .0.10. *Seja A um conjunto não vazio. Uma partição de um conjunto A é uma coleção de subconjuntos não vazios $\{A_1, A_2, \dots, A_n\}$ que satisfaz as seguintes propriedades:*

i) $A = A_1 \cup A_2 \cup \dots \cup A_n$;

ii) A_1, A_2, \dots, A_n são mutuamente disjuntos.