

UMA ABORDAGEM GEOMÉTRICA À TEORIA DOS DELINEAMENTOS EXPERIMENTAIS COM UMA APLICAÇÃO AO PROBLEMA DE ESTIMATIVAS NEGATIVAS DE COMPONENTES DE VARIÂNCIAS

Alexandre da Silva ADÃO¹
Lucas Monteiro CHAVES²
Devanil Jaques de SOUZA²

- RESUMO: Apresentam-se alguns aspectos das teorias de delineamentos experimentais e análise de variância por meio de uma abordagem geométrica em termos de subespaços vetoriais e projetores ortogonais. A apresentação é matematicamente rigorosa e de fácil compreensão. Em razão da concisão e generalidade da linguagem, vários conceitos de análise de variância, por exemplo, são descritos por elementos geométricos. Um exemplo clássico de estimativas negativas de componentes de variância é abordado pelo método.
- PALAVRAS-CHAVE: Abordagem geométrica; delineamento inteiramente casualizado; subespaço vetorial; projetor ortogonal.

1 Introdução

O planejamento adequado de um experimento é essencial para que, em uma pesquisa, os dados obtidos permitam análise estatística adequada. Em razão da diversidade de delineamentos experimentais possíveis, é necessária toda uma teoria que agrupe e classifique tais delineamentos. Uma linguagem matemática adequada a tal objetivo é a que utiliza a teoria da álgebra linear, em razão de sua natureza dupla: álgebra, com toda a sua abstração e elegância, e geometria, com todo o seu apelo intuitivo. Vetores, ângulos e subespaços vetoriais são conceitos geométricos de fácil visualização. O enfoque geométrico da estatística permite que, com a utilização de alguns princípios básicos, se possa desenvolver uma ampla gama de conceitos. Em Bailey (2004, 2008) tal fato fica bastante evidenciado, uma vez que o método geométrico é utilizado de maneira essencial na derivação de resultados estatísticos relativos aos delineamentos experimentais.

Neste trabalho, resultados já apresentados por Bailey (2004, 2008) são demonstrados. A aplicação ao problema clássico de estimativas negativas de componentes de variância é nova.

¹ Universidade Federal de Lavras – UFLA, Programa de Pós-Graduação em Estatística e Experimentação Agropecuária, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: alexandreadao@hotmail.com

² Universidade Federal de Lavras – UFLA, Departamento de Ciências Exatas, Caixa Postal 3037, CEP: 37200-000, Lavras, MG, Brasil. E-mail: lucas@dex.ufla.br / devaniljaques@dex.ufla.br

2 Delineamentos experimentais

Uma unidade experimental, também denominada parcela, é a menor unidade na qual um tratamento pode ser aplicado. O conjunto de todas as unidades experimentais será denominado Ω e as parcelas serão designadas por letras gregas minúsculas $(\alpha, \beta, \dots, \omega, \dots)$. Um tratamento é uma descrição do que pode ser aplicado a uma parcela. Tratamentos serão designados por $\theta_1, \theta_2, \dots, \theta_t$, em que t é o número total de tratamentos. O conjunto de todos os tratamentos será denotado por \mathcal{T} . Um delineamento é uma maneira de se alocarem tratamentos a parcelas, isto é, um delineamento fica definido por uma função $T: \Omega \rightarrow \mathcal{T}$.

Em situações práticas é interessante que as unidades experimentais estejam agrupadas da forma mais homogênea possível. Tais agrupamentos são chamados de blocos. Portanto, a blocagem é uma partição do conjunto Ω das unidades experimentais em b blocos denominados B_1, B_2, \dots, B_b , de modo que $\Omega = \bigcup_{i=1}^b B_i$ e $B_i \cap B_j = \emptyset$ para $i \neq j$. A complexidade combinatória da teoria dos delineamentos é proveniente das relações entre a partição definida pela blocagem e pelas pré-imagens dos tratamentos $T^{-1}(\theta_i) = \{\omega \in \Omega : T(\omega) = \theta_i\}$, $\Omega = \bigcup_{i=1}^t T^{-1}(\theta_i)$.

Note que os subconjuntos $T(B_i)$ não definem uma partição em \mathcal{T} , pois a interseção $T(B_i) \cap T(B_j)$, com $i \neq j$, pode não ser vazia, significando que existem parcelas $\omega \in B_i$ e $\alpha \in B_j$, tais que $T(\omega) = T(\alpha)$, isto é, às parcelas ω e α foi aplicado o mesmo tratamento. Um exemplo simples é o caso do bloco completo, em que a função T , restrita a cada bloco, é uma bijeção sobre \mathcal{T} , significando que em cada bloco estão aplicados todos os tratamentos.

A cardinalidade do conjunto $T^{-1}(\theta_i)$ é denotada por r_i , significando que o tratamento θ_i foi replicado r_i vezes, isto é, aplicado em r_i parcelas. Note que

$$\sum_{i=1}^t r_i = |\Omega| = n, \text{ em que } n \text{ é o número total de parcelas.}$$

2.1 Abordagem geométrica

Dado um delineamento $T: \Omega \rightarrow \mathcal{T}$, como se pode relacionar o delineamento com o espaço \mathbb{R}^n ?

A ideia é: primeiramente ordena-se o conjunto de parcelas de maneira conveniente e associa-se a resposta em cada parcela à coordenada correspondente de uma n -upla no espaço vetorial $V = \mathbb{R}^{|\Omega|} = \mathbb{R}^n$.

Considere o subespaço $V_T \subset V$, denominado subespaço dos tratamentos, formado pelos vetores cujas coordenadas relativas às parcelas que receberam o mesmo tratamento são iguais. Neste caso, é fácil ver que $\dim V_T = t$. Para cada tratamento θ_i , seja u_i o vetor de V cujas coordenadas correspondentes às parcelas que receberam o tratamento θ_i são iguais a 1 e iguais a 0 nas demais coordenadas. Observe que o conjunto $\{u_1, u_2, \dots, u_t\}$ forma uma base ortogonal para V_T .

Seja V_0 o subespaço unidimensional gerado pelo vetor $u_0 = (1, 1, 1, \dots, 1)$. Note que, se $w = (w_1, w_2, \dots, w_n)$ é ortogonal ao subespaço V_0 , então $\sum_{j=1}^n w_j = 0$.

Seja o vetor $Y = (Y_1, \dots, Y_n)$ em que Y_i é a variável aleatória resposta obtida na i -ésima parcela. Sendo assim, Y é um vetor aleatório no espaço vetorial V .

O subespaço vetorial dos blocos, denotado por V_B , é definido como todos os vetores de V cujas coordenadas correspondentes às parcelas em um mesmo bloco são iguais. Seja v_j o vetor de V cujas coordenadas correspondentes a parcelas que pertencem ao bloco B_j são iguais a 1 e iguais a 0 nas demais coordenadas. Então o conjunto $\{v_1, v_2, \dots, v_b\}$ é uma base ortogonal de V_B , ou seja, o produto interno $v_j \cdot v_j = k$, $v_j \cdot v_l = 0$ se $j \neq l$ e $\dim V_B = b$. Note que $u_0 = \sum_{j=1}^b v_j \in V_B$.

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

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

Um delineamento em blocos é dito ortogonal se os subespaços W_T e W_B são ortogonais entre si (Figura 1).

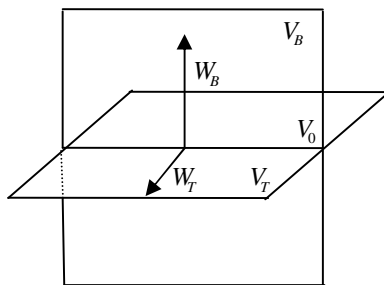


Figura 1 - Representação de um delineamento ortogonal.

A propriedade de ortogonalidade de delineamentos determina certa uniformidade do delineamento.

2.2 Vetores aleatórios

Para modelar geometricamente as propriedades do delineamento tem-se que analisar o comportamento dos vetores obtidos pela projeção do vetor de observações $Y = (Y_1, \dots, Y_n)$ nos subespaços V_0, V_T, V_B, W_T, W_B e no subespaço ortogonal a V_T , denominado V_T^\perp .

A projeção de Y em V_0 :

$$\begin{aligned} P_{V_0} Y &= \left(\frac{Y \cdot u_0}{u_0 \cdot u_0} \right) u_0 = \left(\frac{Y \cdot u_0}{n} \right) u_0 = \frac{1}{n} ((Y_1, \dots, Y_n) \cdot (1, \dots, 1)) (1, \dots, 1) \\ &= \frac{1}{n} (\sum Y_i) (1, \dots, 1) = \frac{1}{n} (\sum Y_1, \sum Y_2, \dots, \sum Y_n) \\ &= \left(\frac{\sum Y_i}{n}, \dots, \frac{\sum Y_i}{n} \right) = (\bar{Y}, \bar{Y}, \dots, \bar{Y}) = \bar{Y} u_0. \end{aligned}$$

$$\text{Portanto, } \|P_{V_0} Y\|^2 = \sum_{i=1}^n (\bar{Y})^2 = n(\bar{Y})^2.$$

A projeção de Y em V_T :

$$P_{V_T} Y = \left(\frac{Y \cdot u_1}{u_1 \cdot u_1} \right) u_1 + \dots + \left(\frac{Y \cdot u_t}{u_t \cdot u_t} \right) u_t.$$

Observe que $Y \cdot u_i$ é a soma de todas as respostas nas parcelas que receberam o tratamento θ_i e $u_i \cdot u_i = r_i$ o número de parcelas que receberam o tratamento θ_i .

$$\|P_{V_T} Y\|^2 = \sum_{i=1}^t \left(\frac{Y \cdot u_i}{r_i} \right)^2 \|u_i\|^2 = \sum_{i=1}^t \left(\frac{Y \cdot u_i}{r_i} \right)^2 r_i = \sum_{i=1}^t \frac{(Y \cdot u_i)^2}{r_i}.$$

Como qualquer vetor em W_T tem soma de coordenadas igual a 0, uma base para o subespaço W_T não seria muito adequada para os cálculos. A ideia, então, é trabalhar com a diferença entre projeções, isto é, $P_{W_T} Y = P_{V_T} Y - P_{V_0} Y$, pois:

$$\begin{aligned} u_0 \cdot P_{W_T} Y &= u_0 \cdot (P_{V_T} Y - P_{V_0} Y) \\ &= u_0 \cdot \left(\left(\sum_1^t \frac{Y \cdot u_i}{u_i \cdot u_i} u_i \right) - \bar{Y} u_0 \right) = \left(\sum_1^t \frac{Y \cdot u_i}{u_i \cdot u_i} u_i \cdot u_0 \right) - \bar{Y} u_0 \cdot u_0 \end{aligned}$$

$$\begin{aligned}
&= \left(\sum_1^t \frac{Y.u_i}{r_i} \right) - \bar{Y}n = \left(\sum_1^t Y.u_i \right) - \bar{Y}n \\
&= \left(\sum_1^t Y_i \right) - \bar{Y}n = \bar{Y}n - \bar{Y}n = 0.
\end{aligned}$$

Da mesma forma, a projeção no V_T^\perp é obtida por diferença, isto é, $P_{V_T^\perp} Y = Y - P_{V_T} Y$.

Observe que projetar o vetor de observações Y no subespaço V_T equivale ao cálculo das médias de cada tratamento. Na ausência de aleatoriedade, as parcelas que receberam o mesmo tratamento deveriam ter a mesma resposta e, portanto, o vetor de observações Y deveria estar contido em V_T . Nesse sentido, pode-se denominar $P_{V_T} Y$ de vetor dos valores ajustados e $P_{V_T^\perp} Y$ de vetor dos resíduos, sendo que este último representa o quanto os dados se afastaram do subespaço de tratamentos em razão dos erros.

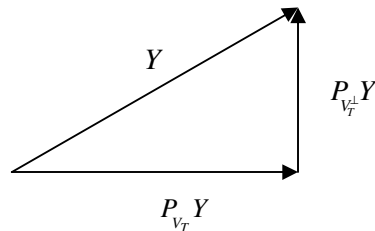


Figura 2 - As projeções $P_{V_T} Y$ e $P_{V_T^\perp} Y$.

Em razão da ortogonalidade (Figura 2) vale o teorema de Pitágoras:

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

As propriedades probabilísticas do vetor aleatório Y , condensadas na proposição 1 abaixo, estão em Bailey (2008), sem demonstrações.

Proposição 1: Seja Y um vetor aleatório n -dimensional, com $E[Y] = \tau$ e $\text{var}[Y] = \Sigma$. Sejam x e z autovetores relativos, respectivamente, aos autovalores ξ e η de Σ . Então:

- i) $\text{var}[x \cdot Y] = \|x\|^2 \xi$.
- ii) Se $\xi = \eta$, $\text{cov}[x \cdot Y, z \cdot Y] = (x \cdot z)\eta$.
Se $\xi \neq \eta$, $\text{cov}[x \cdot Y, z \cdot Y] = 0$.

iii) Se W é um subespaço d -dimensional do subespaço dos autovetores relativos ao autovalor ξ , então:

$$E\left[\|P_W Y\|^2\right] = \|P_W \tau\|^2 + d\xi.$$

Se o vetor aleatório Y tem distribuição normal e W é um subespaço vetorial d -dimensional, consistindo de autovetores de Σ relativos ao autovalor ξ , então:

iv) $\frac{\|P_W(Y - \tau)\|^2}{\xi}$ tem distribuição qui-quadrado com d graus de liberdade.

v) Se $v \in W^\perp$, $\sqrt{d} \frac{x \cdot Y - v \cdot \tau}{\|v\| \|P_W(Y - \tau)\|}$ tem distribuição t de Student com d graus de liberdade.

vi) Se W_1 e W_2 são subespaços de W , ortogonais entre si, com dimensões d_1 e d_2 ,

respectivamente, então $\frac{\|P_{W_1}(Y - \tau)\|^2}{d_1} / \frac{\|P_{W_2}(Y - \tau)\|^2}{d_2}$ tem distribuição F com d_1 e d_2 graus de liberdade.

Demonstrações:

i) $\text{var}[x \cdot Y] = x' \Sigma x = x' (\xi x) = (x \cdot x) \xi = \|x\|^2 \xi.$

ii) $\text{cov}[x \cdot Y, z \cdot Y] = x' \Sigma z = x' \xi z = (x \cdot z) \eta = 0.$

Se $\xi \neq \eta$ então, como os autovetores correspondentes a autovalores diferentes são ortogonais, $\text{cov}[x \cdot Y, z \cdot Y] = (x \cdot z) \eta = 0.$

iii) Seja $\{w_1, \dots, w_d\}$ uma base ortonormal de W . Então:

$$\begin{aligned} P_W Y &= (w_1 \cdot Y) w_1 + \dots + (w_d \cdot Y) w_d \\ E\left[\|P_W Y\|^2\right] &= E\left[(w_1 \cdot Y)^2\right] + \dots + E\left[(w_d \cdot Y)^2\right] \\ &= \text{var}[w_1 \cdot Y] + E[w_1 \cdot Y]^2 + \dots + \text{var}[w_d \cdot Y] + E[w_d \cdot Y]^2 \\ &= w_1' \Sigma w_1 + \dots + w_d' \Sigma w_d + (w_1 \cdot \tau)^2 + \dots + (w_d \cdot \tau)^2 \\ &= \xi \|w_1\|^2 + \dots + \xi \|w_d\|^2 + (w_1 \cdot \tau)^2 + \dots + (w_d \cdot \tau)^2 \\ &= d\xi + \|P_W \tau\|^2. \end{aligned}$$

iv) Temos que $P_w Y = (w_1 \cdot Y)w_1 + \dots + (w_d \cdot Y)w_d$, $E[w_i \cdot Y] = w_i \cdot \tau$ e $\text{var}[w_i \cdot Y] = w_i' \Sigma w_i = \xi$ e, portanto, $w_i \cdot Y$ são variáveis normais independentes de média $w_i \cdot \tau$ e variância ξ . Segue que

$$\begin{aligned} \frac{\|P_w(Y - \tau)\|^2}{\xi} &= \frac{(w_1 \cdot (Y - \tau))^2 + \dots + (w_d \cdot (Y - \tau))^2}{\xi} \\ &= \left(\frac{w_1 \cdot Y - w_1 \cdot \tau}{\sqrt{\xi}} \right)^2 + \dots + \left(\frac{w_d \cdot Y - w_d \cdot \tau}{\sqrt{\xi}} \right)^2 \end{aligned}$$

é uma qui-quadrado com d graus de liberdade.

v) $x \cdot Y$ é uma variável normal com $E[x \cdot Y] = x \cdot \tau$ e $\text{var}[x \cdot Y] = x' \Sigma x = x' \xi x = \|x\|^2 \xi$, portanto $\frac{x \cdot Y - x \cdot \tau}{\|x\| \cdot \sqrt{\xi}}$ tem distribuição normal

padrão independente de $\frac{\|P_w(Y - \tau)\|^2}{\xi}$. Assim, a razão

$$\frac{x \cdot Y - x \cdot \tau}{\|x\| \cdot \sqrt{\xi}} \bigg/ \sqrt{\frac{\|P_w(Y - \tau)\|^2}{\xi} / d} = \sqrt{d} \frac{x \cdot Y - x \cdot \tau}{\|x\| \|P_w(Y - \tau)\|}$$

tem distribuição t de Student com d graus de liberdade.

vi) Como W_1 e W_2 são subespaços ortogonais $P_{W_1} Y$ e $P_{W_2} Y$ são vetores aleatórios

independentes e $\frac{\|P_{W_1}(Y - \tau)\|^2}{\xi}$, $\frac{\|P_{W_2}(Y - \tau)\|^2}{\xi}$ têm distribuições qui-quadrado

com, respectivamente, d_1 e d_2 graus de liberdade, e, portanto, o quociente

$$\frac{\|P_{W_1}(Y - \tau)\|^2}{d_1} \bigg/ \frac{\|P_{W_2}(Y - \tau)\|^2}{d_2}$$

tem distribuição F com parâmetros d_1 e d_2 .

3 Análise de variância

A análise de variância de um delineamento não estruturado está baseada na decomposição do espaço vetorial V como soma dos três subespaços vetoriais ortogonais

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

Os conceitos da análise de variância, em termos geométricos, são apresentados na Tabela 1.

Tabela 1 - Elementos de análise de variância em termos das projeções

Soma de quadrados de tratamentos	$\ P_{W_T} y\ ^2$
Soma de quadrados do resíduo	$\ y - P_{V_T} y\ ^2$
Soma de quadrados total	$\ y - P_{V_0} y\ ^2$
Graus de liberdade para tratamentos	$dim W_T = dim V_T - dim V_0$
Graus de liberdade do resíduo	$dim \mathbb{R}^n - dim V_T$
Grau de liberdade total	$dim \mathbb{R}^n - dim V_0$
Quadrado médio de tratamentos	$\ P_{W_T} y\ ^2 / dim W_T$
Quadrado médio do resíduo	$\ y - P_{V_T} y\ ^2 / (dim \mathbb{R}^n - dim V_T)$

Tabela 2 - Análise de variância de um experimento inteiramente casualizado

	V	V_0	W_T	V_T^\perp
Dimensão	n	1	$(t-1)$	$(n-t)$
Vetor de dados	y	$P_{V_0} y$	$P_{V_T} y - P_{V_0} y$	$(y - P_{V_T} y)$
Soma de quadrados	$\ y\ ^2$	$\frac{\ P_{V_0} y\ ^2}{n}$	$\ P_{V_T} y - P_{V_0} y\ ^2$	$\ (y - P_{V_T} y)\ ^2$
Quadrado médio		$\frac{\ P_{V_0} y\ ^2}{n}$	$\frac{\ P_{V_T} y - P_{V_0} y\ ^2}{t-1}$	$\frac{\ (y - P_{V_T} y)\ ^2}{n-t}$
Esperança do quadrado médio		$n\tau^{-2} + \sigma^2$	$\frac{\sum_i r_i \tau_i^2 - n\tau^{-2}}{t-1} + \sigma^2$	σ^2

Quando se utiliza blocagem, a análise de variância fica definida pela decomposição:

$$V = V_0 \oplus W_T \oplus W_B \oplus V_{(T+B)}^\perp.$$

4 Exemplos

4.1 Aplicação dos conceitos geométricos e análise de variância

Como um exemplo didático para ilustrar a aplicação dos conceitos geométricos na análise de variância, foram considerados dados provenientes de um experimento

conduzido em delineamento inteiramente casualizado, no qual foi avaliado o peso por fruto de determinada variedade de tomates, submetida a três diferentes tipos de fertilizantes. Os dados da Tabela 3 foram adaptados de Searle (1987).

Tabela 3 - Peso por fruto de determinada variedade de tomates em função de três diferentes tipos de fertilizantes ou tratamentos

Tratamentos	Repetições		
	1	2	3
1	74	68	77
2	76	80	
3	87	91	

Seguem os vetores associados:

Vetor resposta: $y = (74, 68, 77, 76, 80, 87, 91)$.

Uma base ortogonal para o subespaço de tratamentos V_T :

$u_1 = (1, 1, 1, 0, 0, 0, 0)$; $u_2 = (0, 0, 0, 1, 1, 0, 0)$; $u_3 = (0, 0, 0, 0, 0, 1, 1)$.

A projeção $P_{V_T} y$:

$$\begin{aligned}
 P_{V_T} y &= \left(\frac{y \cdot u_1}{u_1 \cdot u_1} \right) u_1 + \left(\frac{y \cdot u_2}{u_2 \cdot u_2} \right) u_2 + \left(\frac{y \cdot u_3}{u_3 \cdot u_3} \right) u_3 \\
 &= \left(\frac{(74, 68, 77, 76, 80, 87, 91) \cdot (1, 1, 1, 0, 0, 0, 0)}{(1, 1, 1, 0, 0, 0, 0) \cdot (1, 1, 1, 0, 0, 0, 0)} \right) (1, 1, 1, 0, 0, 0, 0) + \\
 &\quad \left(\frac{(74, 68, 77, 76, 80, 87, 91) \cdot (0, 0, 0, 1, 1, 0, 0)}{(0, 0, 0, 1, 1, 0, 0) \cdot (0, 0, 0, 1, 1, 0, 0)} \right) (0, 0, 0, 1, 1, 0, 0) + \\
 &\quad \left(\frac{(74, 68, 77, 76, 80, 87, 91) \cdot (0, 0, 0, 0, 0, 1, 1)}{(0, 0, 0, 0, 0, 1, 1) \cdot (0, 0, 0, 0, 0, 1, 1)} \right) (0, 0, 0, 0, 0, 1, 1) \\
 &= (73, 73, 73, 78, 78, 89, 89).
 \end{aligned}$$

O vetor de resíduos:

$$\begin{aligned}
 y - P_{V_T} y &= (74 - 73, 68 - 73, 77 - 73, 76 - 78, 80 - 78, 87 - 89, 91 - 89) \\
 &= (1, -5, 4, -2, 2, -2, 2).
 \end{aligned}$$

O vetor de média geral:

$$\begin{aligned}
 P_{V_0} y &= \left(\frac{(74, 68, 77, 76, 80, 87, 91) \cdot (1, 1, 1, 1, 1, 1, 1)}{(1, 1, 1, 1, 1, 1, 1) \cdot (1, 1, 1, 1, 1, 1, 1)} \right) (1, 1, 1, 1, 1, 1, 1) \\
 &= (79, 79, 79, 79, 79, 79, 79).
 \end{aligned}$$

O vetor de observações corrigido pela média geral:

$$y - P_{V_0} y = (-5, -11, -2, -3, 1, 8, 12).$$

O vetor dos contrastes de tratamento:

$$\begin{aligned} P_{W_T} y &= P_{V_T} y - P_{V_0} y \\ &= (73 - 79, 73 - 79, 73 - 79, 78 - 79, 78 - 79, 89 - 79, 89 - 79) \\ &= (-6, -6, -6, -1, -1, 10, 10). \end{aligned}$$

A soma de quadrados de tratamentos:

$$\begin{aligned} \|P_{W_T} y\|^2 &= (36 + 36 + 36 + 1 + 1 + 100 + 100) = 310 \\ \dim W_T &= \dim V_T - \dim V_0 = 3 - 1 = 2. \end{aligned}$$

A soma de quadrados de resíduos (SQR):

$$\begin{aligned} \|y - P_{V_T} y\|^2 &= (1 + 25 + 16 + 4 + 4 + 2 + 9) = 58 \\ \dim SQR &= \dim \mathbb{R}^7 - \dim V_T = 7 - 3 = 4. \end{aligned}$$

A soma de quadrados total (SQT):

$$\begin{aligned} \|y - P_{V_0} y\|^2 &= (25 + 121 + 4 + 9 + 1 + 81 + 144) = 368 \\ \dim SQT &= \dim \mathbb{R}^7 - \dim V_0 = 7 - 1 = 6. \end{aligned}$$

Tabela 4 - Análise de variância

	W_T	V_T^\perp
Dimensão	2	4
Soma de quadrados	310	58
Quadrado médio	155	14,5

4.2 Estimativas de componentes de variância negativos

Uma das técnicas usuais de estimação dos componentes de variância é o método dos momentos, que consiste em igualar os quadrados médios às suas esperanças. Frequentemente, tal método leva a estimativas negativas que, evidentemente, são inaceitáveis. Para explicitar o aspecto geométrico de tais situações vejamos o exemplo discutido em Thompson (1962). Considere que, em K noites, J observações astronômicas de uma mesma quantidade são tomadas por noite. Considerando o efeito “noite” como aleatório, temos o modelo:

$$Y_{kj} = \mu + a_k + e_{kj},$$

em que μ é o verdadeiro valor da medida, e as variáveis aleatórias a_k e e_{kj} são independentes e normais com média zero e variâncias σ_a^2 e σ^2 , respectivamente. Tomando noites como os tratamentos e as medidas de uma mesma noite como repetição, o espaço dos tratamentos é K -dimensional dado por $V_T = \left\{ \left(\underbrace{u_1, \dots, u_1}_J, \underbrace{u_2, \dots, u_2}_J, \dots, \underbrace{u_K, \dots, u_K}_J \right), u_i \in \mathbb{R} \right\}$, contido no espaço de dados

$V = \mathbb{R}^{KJ}$. Observe que a dimensão de V_T^\perp é $KJ - K = K(J - 1)$.

A matriz de variâncias e covariâncias de $Y = (Y_{11}, Y_{12}, \dots, Y_{JK})$ é dada por:

$$i) \quad \text{cov}[Y_{pj}, Y_{qs}] = 0 \quad (p \neq q).$$

$$\begin{aligned} ii) \quad \text{cov}[Y_{kj}, Y_{ks}] &= E[Y_{kj}Y_{ks}] - E[Y_{kj}]E[Y_{ks}] \\ &= E[(\mu + a_k + e_{kj})(\mu + a_k + e_{ks})] \\ &\quad - E[\mu + a_k + e_{kj}]E[\mu + a_k + e_{ks}] \\ &= E[\mu^2 + 2\mu a_k + \mu e_{ks} + a_k^2 + a_k e_{ks} \\ &\quad + \mu e_{kj} + a_k e_{kj} + e_{kj} e_{ks}] - \mu^2 \\ &= \mu^2 + E[a_k^2] - \mu^2 \\ &= \sigma_a^2 \quad (j \neq s). \end{aligned}$$

$$\begin{aligned} iii) \quad \text{cov}[Y_{kj}, Y_{kj}] &= \text{var}[Y_{kj}] = \text{var}[\mu + a_k + e_{kj}] \\ &= \text{var}[a_k] + \text{var}[e_{kj}] \\ &= \sigma_a^2 + \sigma^2. \end{aligned}$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_K \end{bmatrix}$$

$$\Sigma_i = \begin{bmatrix} \sigma_a^2 + \sigma^2 & \sigma_a^2 & \dots & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \sigma^2 & \dots & \sigma_a^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_a^2 & \sigma_a^2 & \dots & \sigma_a^2 + \sigma^2 \end{bmatrix} \quad (i = 1, \dots, K).$$

Isto é, $\Sigma_i = \sigma_a^2 M + \sigma^2 I$, em que M é uma matriz com todas as entradas iguais a 1 e I a matriz identidade, ambas de tamanho $J \times J$.

Os autovalores de Σ são $J\sigma_a^2 + \sigma^2$, relativo ao autoespaço unidimensional $V_0 = \{(u, u, \dots, u), u \in \mathbb{R}\}$, e σ^2 , relativo ao autoespaço $W = \left\{ (u_1, u_2, \dots, u_{KJ}), u_i \in \mathbb{R}, \sum_{i=1}^{KJ} u_i = 0 \right\}$, de dimensão $KJ - 1$. Observe que $V_T^\perp \subset W$. A projeção ortogonal do vetor de dados Y em V_T é dada por $P_{V_T} Y = \bar{Y}_K = \left(\underbrace{\bar{Y}_{1\cdot}, \dots, \bar{Y}_{1\cdot}}_J, \underbrace{\bar{Y}_{2\cdot}, \dots, \bar{Y}_{2\cdot}}_J, \dots, \underbrace{\bar{Y}_{K\cdot}, \dots, \bar{Y}_{K\cdot}}_J \right)$, em que $\bar{Y}_{i\cdot} = \frac{1}{J} \sum_{j=1}^J Y_{ij}$. Logo, $Y - \bar{Y}_K = P_{V_T^\perp} Y$.

Conforme o item *iii* da proposição 1,

$$E \left[\left\| P_{V_T^\perp} Y \right\|^2 \right] = (KJ - K) \sigma^2$$

A matriz de covariâncias de $\bar{Y}_K = \left(\underbrace{\bar{Y}_{1\cdot}, \dots, \bar{Y}_{1\cdot}}_J, \underbrace{\bar{Y}_{2\cdot}, \dots, \bar{Y}_{2\cdot}}_J, \dots, \underbrace{\bar{Y}_{K\cdot}, \dots, \bar{Y}_{K\cdot}}_J \right)$

é dada por:

$$i) \quad \text{cov}[\bar{Y}_{i\cdot}, \bar{Y}_{j\cdot}] = 0 \quad (i \neq j).$$

$$ii) \quad \text{cov}[\bar{Y}_{i\cdot}, \bar{Y}_{i\cdot}] = \text{var}[\bar{Y}_{i\cdot}] = \text{var} \left[\frac{1}{J} \sum_{j=1}^J (\mu + a_i + e_{ij}) \right] \\ = \sigma_a^2 + \frac{1}{J^2} \sum_{j=1}^J \text{var}[e_{ij}] = \sigma_a^2 + \frac{1}{J} \sigma^2$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 & \dots & 0 \\ 0 & \Sigma_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \Sigma_K \end{bmatrix}$$

$$\Sigma_i = \begin{bmatrix} \sigma_a^2 + \frac{\sigma^2}{J} & \sigma_a^2 + \frac{\sigma^2}{J} & \dots & \sigma_a^2 + \frac{\sigma^2}{J} \\ \sigma_a^2 + \frac{\sigma^2}{J} & \sigma_a^2 + \frac{\sigma^2}{J} & \dots & \sigma_a^2 + \frac{\sigma^2}{J} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_a^2 + \frac{\sigma^2}{J} & \sigma_a^2 + \frac{\sigma^2}{J} & \dots & \sigma_a^2 + \frac{\sigma^2}{J} \end{bmatrix}_{(J \times J)} \quad (i = 1, \dots, K).$$

Os autovalores de Σ são $J\sigma_a^2 + \sigma^2$ e 0, relativos, respectivamente, aos autoespaços V_T e V_T^\perp . O vetor $\bar{Y}_K - \bar{Y}_{..}$ é a projeção do vetor \bar{Y}_K no subespaço W_T de V_T , isto é, $P_{W_T} \bar{Y}_K = \bar{Y}_K - \bar{Y}_{..}$. Logo, $\|P_{W_T} \bar{Y}_K\|^2 = \|\bar{Y}_K - \bar{Y}_{..}\|^2 = J \sum_{k=1}^K (\bar{Y}_{k\cdot} - \bar{Y}_{..})^2$. Por *iii* da proposição 1, $E\left[\|P_{W_T} \bar{Y}_K\|^2\right] = \|P_{W_T} \mu\|^2 + (J\sigma_a^2 + \sigma^2)(K-1) = (J\sigma_a^2 + \sigma^2)(K-1)$. Portanto,

a esperança do quadrado médio é dada por $\frac{E\left[\|P_{W_T} \bar{Y}_K\|^2\right]}{(K-1)} = J\sigma_a^2 + \sigma^2$.

Segue então que, pelo método dos momentos, uma estimativa de σ_a^2 é dada por $\frac{1}{J} \left(\frac{\|\bar{Y}_K - \bar{Y}_{..}\|^2}{K-1} - \frac{\|Y - \bar{Y}_K\|^2}{KJ - K} \right)$ e, eventualmente, pode resultar em um valor negativo.

Observe que a diferença entre as somas de quadrados $\|\bar{Y}_K - \bar{Y}_{..}\|^2 - \|Y - \bar{Y}_K\|^2$ pode ser vista como a diferença dos quadrados de dois catetos, conforme Figura 3.

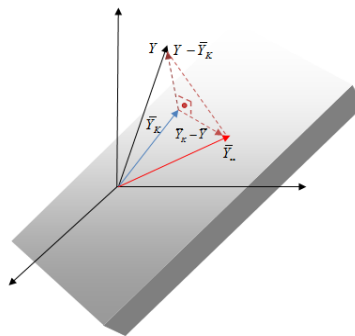


Figura 3 - Diferença de catetos.

Pela proposição 1, a soma de quadrados $\frac{\|\bar{Y}_K - \bar{Y}\|^2}{J\sigma_a^2 + \sigma^2}$ tem distribuição qui-
quadrado com $K - 1$ graus de liberdade. Logo, $\|\bar{Y}_K - \bar{Y}\|^2$ tem distribuição gama com
parâmetros $\frac{J\sigma_a^2 + \sigma^2}{2}$ e $\frac{K - 1}{2}$.

Pela proposição 1, a soma de quadrados $\frac{\|Y - \bar{Y}_K\|^2}{\sigma^2}$ tem distribuição qui-
quadrado com $K(J - 1)$ graus de liberdade. Logo, $\|\bar{Y}_K - \bar{Y}\|^2$ tem distribuição gama
com parâmetros $\frac{\sigma^2}{2}$ e $\frac{K(J - 1)}{2}$.

Os autores consideram que esta abordagem geométrica a este problema clássico de
estimativas negativas para componentes de variância, é original.

Conclusões

- a) A abordagem geométrica dos delineamentos experimentais é didática e conceitualmente interessante;
- b) A soma de quadrados vista como norma de projeções de vetores é concisa e evita fórmulas desnecessariamente longas;
- c) A abordagem geométrica unifica a análise de variância, uma vez que esta depende inteiramente da decomposição do espaço vetorial associado ao conjunto de *parcelas* em termos de subespaços relacionados a tratamentos, blocos e fatores;
- d) A abordagem geométrica permite uma melhor compreensão do problema de estimativas negativas de componentes de variância pelo método dos momentos.

ADÃO, A. S.; CHAVES, L. M.; SOUZA, D. J. Introduction to geometric theory of experimental designs. *Rev. Bras. Biom.*, São Paulo, v.29, n.4, p.628-642, 2011.

- **ABSTRACT:** *We present some aspects of the theory of experimental design and analysis of variance, using a geometric approach in terms of vector subspaces and orthogonal projectors. It is mathematically rigorous and easy to understand. Its main properties are demonstrated using this geometric method. Because of the brevity of geometric language, key concepts as, for example, variance analysis, as described by geometrical elements. A classical example of negative estimates of variance components is treated by the method.*
- **KEYWORDS:** *Geometric approach; completely randomized design; vector subspace; orthogonal projector.*

Referências

- BAILEY, R. A. *Association schemes: designed experiments, algebra and combinatorics*. Cambridge: Cambridge University, 2004. 387p.
- BAILEY, R. A. *Design of comparative experiments*. Cambridge: Cambridge University, 2008. 255p.
- BOLDRINI, J. L.; COSTA, S. I. R.; FIGUEIREDO, V. L.; WETZLER, H. G. *Álgebra linear*. 3.ed. São Paulo: Harper & Row Brasil, 1984. 411p.
- LIMA, E. L. *Álgebra linear*. 7.ed. Rio de Janeiro: IMPA, 2006. 357p.
- PIMENTEL GOMES, F. *Curso de estatística experimental*. 15.ed. Piracicaba: FEALQ, 2009. 467p.
- R DEVELOPMENT CORE TEAM. *R: a language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing, 2009. Disponível em: <<http://www.r-project.org>>. Acesso em: 1 mar. 2010.
- RAO, C. R. *Linear statistical inference and its applications*. New York: J. Wiley, 1973. 626p.
- SEARLE, S. R. *Linear models for unbalanced data*. New York: J. Wiley & Sons, 1987. 536p.
- SIMÕES, J. W. *Métodos de produção de mudas de eucalipto*. Piracicaba: IPEF, 1970. 116p.
- THOMPSON JR., W. A. The problem of negative estimates of variance components. *Ann. Math. Stat.*, Bethesda, v.33, n.1, p.273-289, 1962.

Recebido em 14.08.2011.

Aprovado após revisão 16.02.2012.