Ajuste de Modelos para Experimentos com Medidas Repetidas no R

Fernando Henrique Ferraz P. da Rosa

26 de novembro de 2004

Resumo

Apresentação MAE0327 - Planejamento de Experimentos II, sobre o ajuste de modelos com medidas repetidas no R através do uso de modelos mistos com o pacote lme.

1. Introdução

Há várias abordagens para a modelagem de dados provenientes de experimentos com medidas repetidas. Seja qual for a modelagem adotada, é essencial levar em conta a estrutura de correlação imposta pelo dealineamento. Uma das possíveis maneiras de lidar com isso é através de um modelo linear com efeitos mistos, com efeitos aleatórios para cada bloco de unidade experimental. Essa será a abordagem ilustrada aqui, através do pacote lme[1], do R [2].

2. Modelo para experimentos com um fator e medidas repetidas em todas as unidades experimentais

Consideramos um modelo misto com efeito aleatório para o efeito de unidade experimental:

$$Y_{ij} = \mu + \rho_i + \alpha_j + (\rho\alpha)_{ij} + \epsilon_{ij} \tag{1}$$

onde:

Y_{ij} é o valor da variável resposta para o i-ésimo indivíduo, sob o j-ésimo nível da variável resposta.

μ é uma média global constante

ρ_i é o efeito do i-ésimo indivíduo, seguindo $N(0, \sigma_\rho^2)$.

α_j é o efeito do j-ésimo tratamento, sob a restrição: $\sum_j \alpha_j = 0$.

$(\rho\alpha)_{ij}$ é o efeito de interação indivíduo, tratamento, seguindo $N\left(0, \frac{t-1}{t}\sigma_{\rho\alpha}^2\right)$ restrito a: $\sum_j (\rho\alpha)_{ij} = \sum_i (\rho\alpha)_{ij} = 0$.

ϵ_{ij} são independentes e $N(0, \sigma^2)$

$i = 1, \ldots, n$ **;** $j = 1, \ldots, t$**.**

3. Experimento da Memorização

O experimento foi realizado com o objetivo de se verificar se a memorização imediata de uma sentença por um indivíduo depende da estrutura da sentença. Foi considerado um grupo formado por $n = 10$ indivíduos, com baixa capacidade de memória a curto prazo; os indivíduos ouviam sentanças gravadas em uma fita. Cada sentença era seguida de uma palavra de prova tomada aleatoriamente de uma de $t = 5$ posições. A variável resposta medida foi o tempo de reação de cada indivíduo.

Os dados obtidos foram:

1	2	3	4	5
20	21	42	32	32
67	29	56	39	41
37	25	28	31	34
42	38	36	19	35
57	32	21	30	29
39	38	54	31	28
43	20	46	42	31
35	34	43	35	42
41	23	51	27	30
39	24	35	26	32

Na leitura dos dados no R, especificamos três colunas, uma para reação, outra para posição e outra para sujeito. As 6 primeiras linhas do conjunto de dados nesse formato são:

memorizacao	reacao	posicao	subject
1	20	1	1
2	21	2	1
3	42	3	1
4	32	4	1
5	32	5	1
6	67	1	2

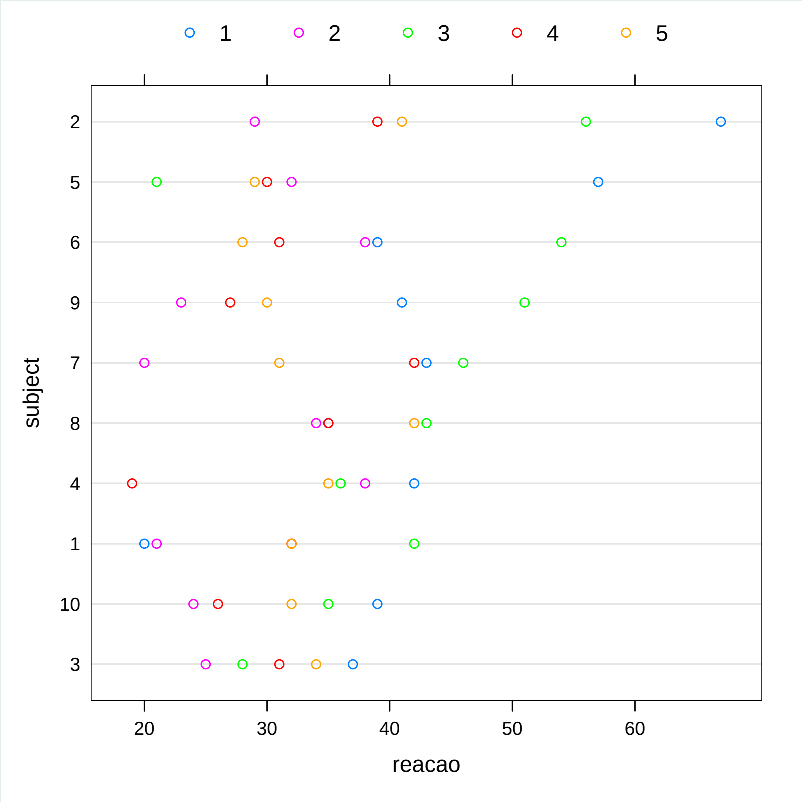
A primeira coisa que fazemos é carregar o pacote `nlme`, através do comando:

```
> library(nlme)
```

Começamos agora criando um outro objeto `memoria`, a partir do conjunto de dados original `memorizacao`, onde armazenamos a estrutura dos dados:

```
> memoria <- groupedData(reacao ~ posicao | subject, data = memorizacao)
```

Esse passo não é obrigatório mas permite facilita a análise, através da criação automática de gráficos por exemplo. Começamos então vendo o gráfico da variável resposta por indivíduo, de acordo com os tratamentos:



Consideramos então o ajuste do modelo **??**. Para fazer o ajuste, rodamos o comando:

```
> z1 <- lme(reacao ~ posicao, data = memoria, random = ~1 | subject/posicao,
```

```
+     method = "ML")
> summary(z1)

Linear mixed-effects model fit by maximum likelihood
 Data: memoria
       AIC      BIC    logLik
  370.3553 385.6514 -177.1776

Random effects:
 Formula: ~1 | subject
        (Intercept)
StdDev:    2.942686

 Formula: ~1 | posicao %in% subject
        (Intercept) Residual
StdDev:    7.454962 2.743107

Fixed effects: reacao ~ posicao
            Value Std.Error DF   t-value p-value
(Intercept) 35.24  1.537661 36 22.917920  0.0000
posicao1     6.76  2.368330 36  2.854332  0.0071
posicao2    -6.84  2.368330 36 -2.888111  0.0065
posicao3     5.96  2.368330 36  2.516541  0.0164
posicao4    -4.04  2.368330 36 -1.705843  0.0967
 Correlation:
         (Intr) posic1 posic2 posic3
posicao1  0.00
posicao2  0.00  -0.25
posicao3  0.00  -0.25  -0.25
```

```
posicao4  0.00  -0.25  -0.25  -0.25

Standardized Within-Group Residuals:
         Min           Q1          Med           Q3          Max
-0.853065699 -0.170434074 -0.008042281  0.140764854  0.889369325

Number of Observations: 50
Number of Groups:
             subject posicao %in% subject
                  10                   50
```

Podemos agora testar os efeitos fixos do modelo, através da tabela de Análise de Vari6ancia:

```
> latex(anova(z1), file = "", table.env = FALSE)
```

anova	numDF	denDF	F-value	p-value
(Intercept)	1	36	525.231070	0.000000000
posicao	4	36	5.266977	0.001913258

Podemos também obter intervalos de confiança para os efeitos do modelo:

```
> intervals(z1)

Approximate 95% confidence intervals

 Fixed effects:
                lower  est.      upper
(Intercept)  32.281511 35.24 38.1984895
```

```
posicao1      2.203289  6.76 11.3167114
posicao2    -11.396711 -6.84 -2.2832886
posicao3      1.403289  5.96 10.5167114
posicao4     -8.596711 -4.04  0.5167114
attr(,"label")
[1] "Fixed effects:"

 Random Effects:
  Level: subject
                   lower     est.    upper
sd((Intercept)) 0.957102 2.942686 9.047522
  Level: posicao
                     lower     est.    upper
sd((Intercept)) 0.01456077 7.454962 3816.863

 Within-group standard error:
       lower          est.         upper
2.773737e-20 2.743107e+00 2.712816e+20
```

Onde confirmamos a suspeita descritiva de que as posições 1 e 3 tem maiores temos de ração, enquanto as posições 2, 4 e 5 menores.

4. Teste de Esfericidade

O teste de esfericidade verifica a propriedade de circularidade da matriz de covariâncias das observações de cada indivíduo. Seja essa matriz de covariâncias dada por Σ_X, e M^* uma matriz $t-1 \times t$ de contrastes ortonormais. Então, de acordo com [3]:

$$M^* \Sigma_X M^{*'} = \lambda I_t$$

Podemos então utilizar o teste de esfericidade de Mauchley, sobre a matriz de covariâncias estimadas ortonormalizadas, para verificar a propriedade de circularide:

Começamos definindo a matriz de contrastes ortonormais:

```
> (M <- t(contr.poly(5)))
```

```
         [,1]       [,2]          [,3]       [,4]      [,5]
.L -0.6324555 -0.3162278 -3.287978e-17  0.3162278 0.6324555
.Q  0.5345225 -0.2672612 -5.345225e-01 -0.2672612 0.5345225
.C -0.3162278  0.6324555  6.329957e-17 -0.6324555 0.3162278
^4  0.1195229 -0.4780914  7.171372e-01 -0.4780914 0.1195229
```

Obtemos então a estimativa de Σ_X:

```
> sig <- var(sapply(levels(posicao), function(x) {
+     reacao[posicao == x]
+ }))
```

E a ortonormalização:

```
> (ort <- M %*% sig %*% t(M))
```

```
           .L         .Q          .C          ^4
.L  76.737778 -47.03565   5.335556   5.257066
.Q -47.035651  81.18730 -11.582370 -36.311614
.C   5.335556 -11.58237  62.373333 -14.898520
^4   5.257066 -36.31161 -14.898520  60.154921
```

Por fim fazemos o teste:

```
> sphericity.test(10, ort)


        Sphericity test

data:  ort
L statistic = 9.53, df = 9, p-value = 0.3899
```

Referências

[1] PINHEIRO, J. C.; BATES, D. M. *Mixed-effects models in S and S-PLUS*. New York: Springer-Verlag, 2000.

[2] R Development Core Team. *R: A language and environment for statistical computing*. Vienna, Austria, 2004. ISBN 3-900051-00-3. Disponível em: <http://www.R-project.org>.

[3] WINER, B. *Statistical Principles in Design*. New York: McGraw-Hill, 1971.

Sobre

A versão eletrônica desse arquivo pode ser obtida em http://www.feferraz.net