MAT02035 - Modelos para dados correlacionados

Modelando a estrutura de covariância

Rodrigo Citton P. dos Reis
citton.padilha@ufrgs.br

Universidade Federal do Rio Grande do Sul
Instituto de Matemática e Estatística
Departamento de Estatística

Porto Alegre, 2023

Introdução

Introdução

- Uma vez que uma das características definidoras dos **dados longitudinais** é que eles **são correlacionados**, devemos considerar abordagens para **modelar** apropriadamente **a covariância** ou dependência do tempo entre as medidas repetidas obtidas nos mesmos indivíduos.
- Quando um modelo apropriado para a covariância é adotado, erros padrão corretos são obtidos e inferências válidas sobre os parâmetros de regressão podem ser feitas.
- Levar em consideração a covariância entre medidas repetidas geralmente aumenta a eficiência ou a precisão com a qual os parâmetros de regressão podem ser estimados; ou seja, a correlação positiva entre as medidas repetidas reduz a variabilidade da estimativa de mudança ao longo do tempo dentro dos indivíduos.

Introdução

- Além disso, quando há dados ausentes, a modelagem correta da covariância é frequentemente um requisito para obter estimativas válidas dos parâmetros de regressão.
- Em geral, a falha em levar em consideração a covariância entre as medidas repetidas resultará em estimativas incorretas da variabilidade da amostragem e pode levar a inferências científicas enganosas.

Introdução

- Os dados longitudinais apresentam **dois aspectos** dos dados que **requerem modelagem**: a **resposta média** ao longo do tempo e a **covariância**.
- Embora esses dois aspectos dos dados possam ser modelados separadamente, eles estão **inter-relacionados**.
- As escolhas de modelos para resposta média e covariância são **interdependentes**.
- Um modelo para a covariância deve ser escolhido com base em algum modelo assumido para a resposta média.
- A covariância entre qualquer par de resíduos, digamos $[Y_{ij} - \mu_{ij}(\beta)]$ e $[Y_{ik} - \mu_{ik}(\beta)]$, depende do modelo para a média condicional.

Introdução

- Um modelo para a covariância deve ser escolhido com base em algum modelo para a resposta média; este representa uma tentativa de levar em consideração a covariância entre os resíduos que resultam de um modelo específico para a média.
- Uma escolha diferente de modelo para a média ou, além disso, qualquer especificação incorreta do modelo para a média, pode potencialmente resultar em uma escolha diferente de modelo para a covariância.
- Como resultado dessa interdependência entre os modelos de média e covariância, precisaremos desenvolver uma estratégia geral de modelagem que leve essa interdependência em consideração.

Introdução

Três abordagens gerais podem ser distinguidas:

(1) padrão de covariância "não estruturado"ou arbitrário;

(2) modelos de padrão de covariância;

(3) estrutura de covariância de efeitos aleatórios.

Covariância não estruturada

Covariância não estruturada

- **Apropriada** quando o **delineamento** é **balanceado** e o **número de ocasiões** de medições é relativamente **pequeno**.
- Nenhuma estrutura explícita é assumida além da **homogeneidade de covariância entre diferentes indivíduos**, $\text{Cov}(Y_i) = \Sigma_i = \Sigma$.
- **Principal vantagem:** nenhuma suposição sobre os padrões de variância e covariâncias.

Covariância não estruturada

- Com n ocasiões de medição, a matriz de covariância "não estruturada" possui $n \times (n+1)/2$ parâmetros:
  - as n variâncias ($\sigma_j^2 = \text{Var}[Y_{ij}]$) e as $n \times (n-1)/2$ covariâncias ($\sigma_{jk} = \text{Cov}[Y_{ij}, Y_{ik}]$) duas-a-duas,

$$\text{Cov}(Y_i) = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \cdots & \sigma_{1n} \\ \sigma_{21} & \sigma_2^2 & \cdots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{n1} & \sigma_{n2} & \cdots & \sigma_n^2 \end{pmatrix}.$$

Covariância não estruturada

Potenciais desvantagens:

- O número de parâmetros de covariância cresce rapidamente com o número de ocasiões de medição:
  - $n = 3$, o número de parâmetros de covariância é 6;
  - $n = 5$, o número de parâmetros de covariância é 15;
  - $n = 10$, o número de parâmetros de covariância é 55.
- Quando o número de parâmetros de covariância é grande, em relação ao tamanho da amostra, é provável que a estimativa seja muito instável.
- O uso de uma covariância não estruturada é atraente apenas quando N é grande em relação a $n \times (n+1)/2$.
- A covariância não estruturada é problemática quando há medições irregulares no tempo.

Modelos de padrão de covariância

Modelos de padrão de covariância

- Ao tentar impor alguma estrutura à covariância, um **equilíbrio** sutil precisa ser alcançado.
  - Com **pouca estrutura**, pode haver **muitos parâmetros** a serem estimados com quantidade limitada de dados.
  - Com **muita estrutura**, risco potencial de **erros de especificação do modelo** e inferências enganosas a respeito de β.
- Este é o clássico **tradeoff** entre **viés** e **variância**.
- Os modelos de padrão de covariância têm como base **modelos de correlação serial** originalmente desenvolvidos para dados de **séries temporais**.
  - A seguir, veremos alguns destes modelos.

Relação entre a covariância e a correlação

- Lembre que $\rho_{jk} = \text{Corr}\left(Y_{ij}, Y_{ik}\right)$ pode ser expresso como

$$\begin{aligned}\rho_{jk} &= \frac{\text{Cov}\left(Y_{ij}, Y_{ik}\right)}{\sqrt{\text{Var}\left(Y_{ij}\right)}\sqrt{\text{Var}\left(Y_{ik}\right)}} \\ &= \frac{\sigma_{jk}}{\sigma_j\sigma_k} \\ &\Rightarrow \sigma_{jk} = \rho_{jk}\sigma_j\sigma_k.\end{aligned}$$

- Se assumirmos que $\text{Var}\left(Y_{i1}\right) = \text{Var}\left(Y_{i2}\right) = \ldots \text{Var}\left(Y_{in}\right) = \sigma^2$, então $\sigma_j = \sigma_k$ e $\sigma_j\sigma_k = \sigma^2$, e assim $\sigma_{jk} = \rho_{jk}\sigma^2$.
  - Ou seja, modelos de padrão de covariância podem ser obtidos a partir da especificação de um padrão da estrutura de correlação.

Simetria composta

- Este padrão assume que a **variância é constante** entre as ocasiões, digamos σ^2, e $\text{Corr}\left(Y_{ij}, Y_{ik}\right) = \rho$ para todo j e k.

$$\text{Cov}\left(Y_i\right) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho & \cdots & \rho \\ \rho & 1 & \rho & \cdots & \rho \\ \rho & \rho & 1 & \cdots & \rho \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho & \rho & \rho & \cdots & 1 \end{pmatrix}.$$

- **Vantagem:** (parcimônia) possui apenas dois parâmetros, independentemente do número de ocasiões de medição.
- **Desvantagem:** Faz fortes suposições sobre variância e correlação que geralmente não são válidas com dados longitudinais.

Toeplitz

- Assume que a variância é constante entre as ocasiões, digamos σ^2, e $\text{Corr}\,(Y_{ij}, Y_{i,j+k}) = \rho_k$ para todo j e k.

$$\text{Cov}\,(Y_i) = \sigma^2 \begin{pmatrix} 1 & \rho_1 & \rho_2 & \cdots & \rho_{n-1} \\ \rho_1 & 1 & \rho_1 & \cdots & \rho_{n-2} \\ \rho_2 & \rho_1 & 1 & \cdots & \rho_{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_{n-1} & \rho_{n-2} & \rho_{n-3} & \cdots & 1 \end{pmatrix}.$$

- Assume que a correlação entre as respostas em ocasiões adjacentes de medição é constante, ρ_1.

Toeplitz

- Toeplitz é apropriado apenas quando as medições são feitas em intervalos de tempo iguais (ou aproximadamente iguais).
- A covariância Toeplitz possui n parâmetros (1 parâmetro de variância e $n - 1$ parâmetros de correlação).
- Um caso especial da covariância Toeplitz é a covariância **autorregressiva** (de primeira ordem).

Autoregressiva

- Assume que a variância é constante entre as ocasiões, digamos σ^2, e $\text{Corr}\,(Y_{ij}, Y_{i,j+k}) = \rho^k$ para todo j e k, e $\rho \geq 0$.

$$\text{Cov}\,(Y_i) = \sigma^2 \begin{pmatrix} 1 & \rho & \rho^2 & \cdots & \rho^{n-1} \\ \rho & 1 & \rho & \cdots & \rho^{n-2} \\ \rho^2 & \rho & 1 & \cdots & \rho^{n-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1} & \rho^{n-2} & \rho^{n-3} & \cdots & 1 \end{pmatrix}.$$

- **Parcimônia:** apenas 2 parâmetros, independentemente do número de ocasiões de medição.
- Somente apropriado quando as medições são feitas em intervalos de tempo iguais (ou aproximadamente iguais).

Autoregressiva

- Simetria composta, Toeplitz e covariância autorregressiva assumem que as variâncias são constantes ao longo do tempo.
- Essa suposição pode ser relaxada considerando-se versões desses modelos com variâncias heterogêneas, $\text{Var}\left(Y_{ij}\right) = \sigma_j^2$.
- Um padrão heterogêneo de covariância autorregressiva:

$$\text{Cov}\left(Y_i\right) = \begin{pmatrix} \sigma_1^2 & \rho\sigma_1\sigma_2 & \rho^2\sigma_1\sigma_3 & \cdots & \rho^{n-1}\sigma_1\sigma_n \\ \rho\sigma_1\sigma_2 & \sigma_2^2 & \rho\sigma_2\sigma_3 & \cdots & \rho^{n-2}\sigma_2\sigma_n \\ \rho^2\sigma_1\sigma_3 & \rho\sigma_2\sigma_3 & \sigma_3^2 & \cdots & \rho^{n-3}\sigma_3\sigma_n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho^{n-1}\sigma_1\sigma_n & \rho^{n-2}\sigma_2\sigma_n & \rho^{n-3}\sigma_3\sigma_n & \cdots & \sigma_n^2 \end{pmatrix}.$$

e possui $n + 1$ parâmetros (n parâmetros de variância e 1 parâmetro de correlação).

Banded (em faixas)

- Assume que a correlação é zero além de algum intervalo especificado.
- Por exemplo, um padrão de covariância em faixas com tamanho de banda 3 pressupõe que $\text{Corr}\left(Y_{ij}, Y_{i,j+k}\right) = 0$ para $k \geq 3$.
- É possível aplicar um padrão em faixas a qualquer um dos modelos de padrão de covariância considerados até o momento.

Banded (em faixas)

- Um padrão de covariância de Toeplitz em faixas com um tamanho de banda 2 é dado por,

$$\mathrm{Cov}\,(Y_i) = \sigma^2 \begin{pmatrix} 1 & \rho_1 & 0 & \cdots & 0 \\ \rho_1 & 1 & \rho_1 & \cdots & 0 \\ 0 & \rho_1 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix}.$$

em que $\rho_2 = \rho_3 = \ldots = \rho_{n-1} = 0$.

- A estrutura banded faz suposições muito fortes sobre a rapidez com que a correlação cai para zero com o aumento da separação entre as ocasiões de medição.

Exponencial

- Quando as ocasiões de medição não são igualmente espaçadas ao longo do tempo, o modelo autorregressivo pode ser generalizado da seguinte maneira.
- Denote os tempos de observação para o i-ésimo indivíduo por $\{t_{i1}, \ldots, t_{in}\}$ e assuma que a variância é constante em todas as ocasiões de medição, digamos σ^2, e

$$\text{Corr}\left(Y_{ij}, Y_{ik}\right) = \rho^{|t_{ij} - t_{ik}|}, \text{ para } \rho \geq 0.$$

- A correlação entre qualquer par de medidas repetidas diminui exponencialmente com as separações de tempo entre elas.

Exponencial

- Este padrão é referido como **covariância "exponencial"** porque pode ser reexpressa como

$$\begin{aligned} \mathrm{Cov}\left(Y_{ij}, Y_{ik}\right) &= \sigma^2 \rho^{|t_{ij} - t_{ik}|} \\ &= \sigma^2 \exp(-\theta |t_{ij} - t_{ik}|), \end{aligned}$$

em que $\theta = -\log(\rho)$ ou $\rho = \exp(-\theta)$ para $\theta \geq 0$.

- O modelo de covariância exponencial é invariante sob transformação linear da escala de tempo.
  - Se substituirmos t_{ij} por $(a + bt_{ij})$ (por exemplo, se substituirmos o tempo medido em "semanas" pelo tempo medido em "dias"), a mesma forma para a matriz de covariância se mantém.

Escolha entre modelos de padrão de covariância

- A escolha de modelos para covariância e média são interdependentes.
- A escolha do modelo de covariância deve ser baseada em um modelo **"maximal"** para a média que minimiza qualquer possível erro de especificação.
- Com delineamentos balanceados e um número muito pequeno de covariáveis discretas, escolha o "modelo saturado" para a resposta média.
- **Relembrando:** o **modelo saturado** inclui os efeitos principais do tempo (considerado como um fator dentro de indivíduo) e todos os outros efeitos principais, além de suas interações duas-a-duas e de ordem superiores.

Escolha entre modelos de padrão de covariância

- De forma mais geral, quando o modelo saturado não puder ser empregado, um modelo maximal deve ser especificado para a média da resposta.
  - O modelo maximal deve ser, em certo sentido, o modelo mais elaborado para a resposta média que consideraríamos do ponto de vista do indivíduo.
- Uma vez escolhido o modelo maximal, a variância e a covariância residual podem ser usadas para selecionar o modelo apropriado para covariância.

Escolha entre modelos de padrão de covariância

- Dado um modelo maximal para a média, uma sequência de modelos de padrão de covariância pode ser ajustada aos dados disponíveis.
- **A escolha** entre os modelos pode ser feita **comparando as verossimilhanças maximizadas** para cada um dos modelos de padrão de covariância.
- Ou seja, quando qualquer par de modelos é encaixado, uma estatística de **teste de razão de verossimilhanças** pode ser construída para comparar os modelos "completo" e "reduzido".

Escolha entre modelos de padrão de covariância

- Lembre-se que dois modelos são encaixados quando o modelo "reduzido" é um caso especial do modelo "completo".
- Por exemplo, o modelo de simetria composta é um caso especial do modelo Toeplitz, desde que $\rho_1 = \rho_2 = \ldots = \rho_{n-1}$.
- A estatística do teste da razão de verossimilhanças é obtido tomando-se o dobro da diferença das respectivas **log-verossimilhanças REML maximizadas**,

$$G^2 = 2(\hat{\ell}_{comp} - \hat{\ell}_{red}),$$

e comparando a estatística com uma distribuição qui-quadrado com graus de liberdade igual à diferença entre o número de parâmetros de covariância nos modelos completo e reduzido.

Escolha entre modelos de padrão de covariância

- **Para comparar modelos não encaixados**, uma abordagem alternativa é o **Critério de Informação de Akaike (Akaike Information Criterion - AIC)**.
- De acordo com o AIC, dado um conjunto de modelos concorrentes para a covariância, **deve-se selecionar o modelo que minimiza**

$$
\begin{aligned}
AIC &= -2(\text{log-vero. maximizada}) + 2(\text{número de parâmetros}) \\
&= -2(\hat{\ell} - c),
\end{aligned}
$$

em que $\hat{\ell}$ é a **log-verossimilhança REML maximizada** e c é o número de parâmetros de covariância.

Exemplo

Experimento de Terapia por Exercício

- Neste estudo, os indivíduos foram designados para um dos dois programas de levantamento de peso para aumentar a força muscular.
    - **Tratamento 1:** o número de repetições dos exercícios foi aumentado à medida que os indivíduos se tornaram mais fortes.
    - **Tratamento 2:** o número de repetições foi mantido constante, mas a quantidade de peso foi aumentada à medida que os indivíduos se tornaram mais fortes.
- As medidas de força corporal foram realizadas na linha de base e nos dias 2, 4, 6, 8, 10 e 12.
- Vamos nos concentrar apenas nas medidas de força obtidas na linha de base (ou no dia 0) e nos dias 4, 6, 8 e 12.

Experimento de Terapia por Exercício

- Antes de considerar modelos para a covariância, é necessário escolher um modelo maximal para a resposta média.
- Neste exemplo (delineamento balanceado e apenas dois grupos de tratamento), optamos pelo modelo saturado como o modelo maximal para a média.
- Primeiramente, vamos considerar uma matriz de **covariância não estruturada**.

Modelo maximal com cov. não estruturada

```
library(nlme)

# matriz de covariância não estruturada
mod1 <- gls(fc ~ trt*dia,
            na.action = na.omit,

            corr = corSymm(form = ~ tempo | id),
            weights = varIdent(form = ~ 1 | tempo),

            method = "REML",
            data = af.longo)
```

Matriz de covariância não estruturada estimada

```
library(lavaSearch2)

knitr::kable(
  getVarCov2(mod1)$Omega,
  digits = 3)
```

1	2	3	4	5
9.668	10.175	8.974	9.812	9.407
10.175	12.550	11.091	12.580	11.928
8.974	11.091	10.642	11.686	11.101
9.812	12.580	11.686	13.991	13.121
9.407	11.928	11.101	13.121	13.945

Matriz de covariância não estruturada estimada

- Observe que a variância parece ser maior no final do estudo, quando comparada à variância na linha de base.
- A correspondente matriz de correlação pode ser obtida com

```
mod1$modelStruct$corStruct
```

```
## Correlation structure of class corSymm representing
##  Correlation:
##   1     2     3     4
## 2 0.924
## 3 0.885 0.960
## 4 0.844 0.949 0.958
## 5 0.810 0.902 0.911 0.939
```

- Note que as correlações diminuem à medida que a separação do tempo entre as medidas repetidas aumenta.

Modelo maximal com cov. autoregressiva

- Apesar do aparente aumento da variância ao longo do tempo, consideramos um **modelo autorregressivo** para a covariância.
- Este modelo é muito parcimonioso, com apenas dois parâmetros, um descrevendo a variância, σ^2, o outro a correlação, ρ.

```
# library(nlme)

# matriz de covariância autorregressiva
mod2 <- gls(fc ~ trt*dia,
            na.action = na.omit,

            corr = corAR1(form = ~ tempo | id),

            method = "REML",
            data = af.longo)
```

Modelo maximal com cov. autoregressiva

- Quando um modelo autorregressivo de primeira ordem é ajustado aos dados, ele resulta nas seguintes estimativas dos parâmetros de variância e correlação

```
summary(mod2)$sigma^2
```

```
## [1] 11.86727
```

```
coef(mod2$modelStruct$corStruct,
     uncons = FALSE, allCoef = TRUE)
```

```
##      Phi1
## 0.9401763
```

Modelo maximal com cov. autoregressiva

- As correlações pareadas estimadas resultantes entre as cinco medições repetidas são fornecidas por

```
getVarCov(mod2)/summary(mod2)$sigma^2
```

```
## Marginal variance covariance matrix
##         [,1]    [,2]    [,3]    [,4]    [,5]
## [1,] 1.00000 0.94018 0.88393 0.83105 0.78133
## [2,] 0.94018 1.00000 0.94018 0.88393 0.83105
## [3,] 0.88393 0.94018 1.00000 0.94018 0.88393
## [4,] 0.83105 0.88393 0.94018 1.00000 0.94018
## [5,] 0.78133 0.83105 0.88393 0.94018 1.00000
##   Standard Deviations: 1 1 1 1 1
```

- O modelo autorregressivo foi ajustado principalmente para fins ilustrativos; o modelo não é muito apropriado para esses dados, pois eles estão espaçados de forma desigual ao longo do tempo (ou seja, há um intervalo de quatro dias entre as duas primeiras medidas repetidas e as duas últimas medidas repetidas, mas todas as outras medidas repetidas adjacentes foram feitas com dois dias de intervalo).

Modelo maximal com cov. exponencial

- A fim de contabilizar o intervalo de tempo desigual, um modelo exponencial para a covariância foi considerado, em que

$$\text{Cov}\left(Y_{ij}, Y_{ik}\right) = \sigma^2 \rho^{|t_{ij} - t_{ik}|},$$

para $t_{i1} = 0,\ t_{i2} = 4,\ t_{i3} = 6,\ t_{i4} = 8$ e $t_{i5} = 12$ para todos os indivíduos.

Modelo maximal com cov. exponencial

```
af.longo$dia.num <- as.numeric(as.character(af.longo$dia))

# library(nlme)

# matriz de covariância exponencial
mod3 <- gls(fc ~ trt*dia,
            na.action = na.omit,

            corr = corExp(form = ~ dia.num | id),

            method = "REML",
            data = af.longo)
```

Modelo maximal com cov. exponencial

- A variância e as correlações pareadas estimadas resultantes entre as cinco medições repetidas são fornecidas por

```
summary(mod3)$sigma^2
```

```
## [1] 11.87491
```

```
getVarCov(mod3)/summary(mod3)$sigma^2
```

```
## Marginal variance covariance matrix
##         [,1]    [,2]    [,3]    [,4]    [,5]
## [1,] 1.00000 0.91694 0.87804 0.84079 0.77096
## [2,] 0.91694 1.00000 0.95757 0.91694 0.84079
## [3,] 0.87804 0.95757 1.00000 0.95757 0.87804
## [4,] 0.84079 0.91694 0.95757 1.00000 0.91694
## [5,] 0.77096 0.84079 0.87804 0.91694 1.00000
##   Standard Deviations: 1 1 1 1 1
```

Modelo maximal com cov. exponencial

- Note que os declínios nas correlações estimadas dos modelos autorregressivo e exponencial são muito rápidos quando comparados aos declínios correspondentes do modelo de covariância não estruturada.

Comparando os modelos

- A seguir, consideramos a escolha entre esses modelos de padrão de covariância.
- A log-verossimilhança REML maximizada e o AIC para cada um dos modelos de padrão de covariância são exibidos a seguir

```
anova(mod1, mod2, mod3, test = FALSE)
```

```
##      Model df      AIC      BIC    logLik
## mod1     1 25 647.3399 724.6837 -298.6699
## mod2     2 12 645.0718 682.1968 -310.5359
## mod3     3 12 642.5459 679.6709 -309.2729
```

Comparando os modelos

- Observe que existe uma hierarquia entre os modelos.
  - Os modelos autorregresivo e exponencial são encaixados no modelo de covariância não estruturada.
- As comparações dos modelos autorregressivo e exponencial com o modelo de covariância não estruturada podem ser feitas usando **testes de razão de verossimilhança (REML)**.
- No entanto, os modelos autorregresivo e exponencial não são modelos encaixados; ainda, ambos os modelos têm o mesmo número de parâmetros.
  - Como resultado, qualquer comparação entre esses dois modelos pode ser feita diretamente em termos de suas log-verossimilhanças maximizadas, uma vez que qualquer penalidade extraída por critérios de informação será a mesma em ambos os casos.

Comparando os modelos

- Com base no teste da razão de verossimilhanças, há evidências de que o modelo autorregressivo não fornece um ajuste adequado para a covariância, quando comparado à covariância não estruturada ($p < 0.05$).

```
anova(mod1, mod2)
```

```
##      Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## mod1     1 25 647.3399 724.6837 -298.6699
## mod2     2 12 645.0718 682.1968 -310.5359 1 vs 2 23.73189  0.0337
```

Comparando os modelos

- Por outro lado, o teste da razão de verossimilhanças, comparando a covariância exponencial e não estruturada, resulta em

```
anova(mod1, mod3)
```

```
##      Model df      AIC      BIC    logLik   Test  L.Ratio p-value
## mod1     1 25 647.3399 724.6837 -298.6699
## mod3     2 12 642.5459 679.6709 -309.2729 1 vs 2 21.20597   0.069
```

- Assim, a covariância exponencial fornece um ajuste adequado aos dados ($p > 0.05$).
- Além disso, em termos do AIC, o modelo exponencial minimiza esse critério.

Pontos fortes e fracos dos modelos de covariância

Pontos fortes e fracos dos modelos de covariância

- Os modelos de padrão de covariância tentam caracterizar a covariância com um número relativamente pequeno de parâmetros.
- No entanto, muitos modelos (por exemplo, autorregresivo, Toeplitz e em faixas) são apropriados apenas quando medições repetidas são obtidas em intervalos iguais e não podem lidar com medições irregulares no tempo.
- Embora exista uma grande variedade de modelos para correlações, a escolha de modelos para variâncias é limitada.
- Eles não são adequados para modelar dados de delineamentos longitudinais inerentemente desbalanceados.

Exercícios

Exercícios

- Resolva os exercícios do Capítulo 7 do livro "**Applied Longitudinal Analysis**" (páginas 186 e 187).
  - O arquivo de dados (`dental.dta`) está no Moodle.

Avisos

Avisos

- **Próxima aula:** Modelos lineares de efeitos mistos.
- **Para casa:** ler o Capítulo 7 do livro "**Applied Longitudinal Analysis**".
  - Caso ainda não tenha lido, leia também os Caps. 1, 2, 3, 4, 5 e 6.

Bons estudos!