

Análise Fatorial: parte 2

Prof. Caio Azevedo

Predição dos fatores

- Uma vez estimadas as cargas fatoriais (\mathbf{L}) e a matriz de variâncias-covariâncias dos fatores específicos (Ψ) (como visto anteriormente) tem-se o interesse em prever os fatores.
- Os valores preditos dos escores, em geral, são chamados de escores fatoriais.
- Veremos dois métodos: mínimos quadrados ponderados e método de regressão.
- Tanto as estimativas de \mathbf{L} e Ψ , obtidas pelo método das componentes principais quanto por máxima verossimilhança, podem ser usadas.

Cont.

- Lembremos que, em geral, se utiliza a matriz de correlações.
- Em geral, também, trabalhamos com as variáveis originais padronizadas (média 0 e variância 1) para gerar os escores.
- Naturalmente, para um mesmo método de predição de escores, os resultados podem ser diferentes consoante o método de estimação usado para as matrizes L e Ψ .
- Os escores fatoriais são usados, em geral, como variáveis resposta (ou explicativas) em análises subsequentes bem como para caracterizar (entender) melhor os dados.

Mínimos quadrados ponderados

- Lembrando que o modelo original é dado por
$$\mathbf{x}_i - \boldsymbol{\mu} = \mathbf{L}\mathbf{f}_i + \boldsymbol{\xi}_i, i = 1, 2, \dots, n.$$
- O objetivo é minimizar a soma de quadrados dos erros ponderada ([link](#)), ou seja, minimizar

$$Q = \sum_{j=1}^p \frac{\xi_{ij}^2}{\psi_j} = \boldsymbol{\xi}_i' \boldsymbol{\Psi}^{-1} \boldsymbol{\xi}_i$$

(para cada indivíduo).

- Temos ainda que $Q = (\mathbf{x}_i - \boldsymbol{\mu} - \mathbf{L}\mathbf{f}_i)' \boldsymbol{\Psi}^{-1} (\mathbf{x}_i - \boldsymbol{\mu} - \mathbf{L}\mathbf{f}_i)$.

Cont.

- Derivando-se Q com relação à f_i , igualando-se o respectivo resultado à zero e resolvendo o sistema resultante, obtém-se

$$\hat{F}_i^* = (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1} (\mathbf{X}_i - \boldsymbol{\mu}).$$

- Substituindo-se Ψ , \mathbf{L} e $\boldsymbol{\mu}$ pelos respectivos estimadores, vem que

$$\hat{F}_i = \left(\hat{\mathbf{L}}' \hat{\Psi}^{-1} \hat{\mathbf{L}} \right)^{-1} \hat{\mathbf{L}}' \hat{\Psi}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}).$$

- Se Ψ e \mathbf{L} tiverem sido estimados pelo método de máxima verossimilhança, então $\mathbf{L}'\Psi^{-1}\mathbf{L} = \Delta$, (se $\hat{\Delta} = \hat{\mathbf{L}}' \hat{\Psi}^{-1} \hat{\mathbf{L}}$), logo

$$\hat{F}_i = \hat{\Delta}^{-1} \hat{\mathbf{L}}' \hat{\Psi}^{-1} (\mathbf{X}_i - \bar{\mathbf{X}}).$$

Cont.

- Note ainda que $\mathcal{E}(\widehat{\mathbf{F}}_i^*) = (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1}(\mathcal{E}(\mathbf{X}_i) - \boldsymbol{\mu}) = \mathbf{0}$.

- Além disso,

$$\begin{aligned} \text{Cov}(\widehat{\mathbf{F}}_i^*) &= (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1} \text{Cov}(\mathbf{X}_i - \boldsymbol{\mu}) \Psi^{-1} \mathbf{L} (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \\ &= (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1} (\mathbf{L}\mathbf{L}' + \Psi) \Psi^{-1} \mathbf{L} (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \\ &= (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1} \mathbf{L}\mathbf{L}'\Psi^{-1} \mathbf{L} (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \\ &\quad + (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \mathbf{L}'\Psi^{-1} \Psi \Psi^{-1} \mathbf{L} (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \\ &= \mathbf{I} + (\mathbf{L}'\Psi^{-1}\mathbf{L})^{-1} \end{aligned}$$

- Se Ψ e \mathbf{L} tiverem sido estimados pelo método de máxima verossimilhança, então $\text{Cov}(\mathbf{F}_i^*) = \mathbf{I} + \Delta^{-1}$. Logo, $\mathcal{E}(\widehat{\mathbf{F}}_i) \approx \mathbf{0}$ e $\text{Cov}(\widehat{\mathbf{F}}_i) \approx \mathbf{I} + \widehat{\Delta}^{-1}$ ou $\left(\mathbf{I} + \left(\widehat{\mathbf{L}}' \widehat{\Psi}^{-1} \widehat{\mathbf{L}} \right)^{-1} \right)$.

Método da regressão

- Supõe-se que, $i=1,2,\dots,n$

$$\begin{bmatrix} \mathbf{X}_i - \boldsymbol{\mu} \\ \mathbf{F}_i \end{bmatrix} \sim N_{(p+m)} \left[\begin{pmatrix} \mathbf{0}_{(p \times 1)} \\ \mathbf{0}_{(m \times 1)} \end{pmatrix}, \begin{pmatrix} \boldsymbol{\Sigma} = \mathbf{L}\mathbf{L}' + \boldsymbol{\Psi} & \mathbf{L} \\ \mathbf{L}' & \mathbf{I} \end{pmatrix} \right]$$

que é equivalente a assumir que $\mathbf{F} \sim N_m(\mathbf{0}, \mathbf{I})$ e $\boldsymbol{\xi} \sim N_p(\mathbf{0}, \boldsymbol{\Psi})$.

- Usando o resultado ([visto anteriormente](#)) sobre distribuições condicionais de normais multivariadas, temos que

$\mathbf{F}|\mathbf{X} \sim N(\boldsymbol{\mu}_F, \boldsymbol{\Psi}_F)$, em que

$$\boldsymbol{\mu}_F = \mathbf{L}'\boldsymbol{\Sigma}^{-1}(\mathbf{x} - \boldsymbol{\mu}) = \mathbf{L}'(\mathbf{L}\mathbf{L}' + \boldsymbol{\Psi})^{-1}(\mathbf{x} - \boldsymbol{\mu})$$

$$\boldsymbol{\Psi}_F = \mathbf{I} - \mathbf{L}'\boldsymbol{\Sigma}^{-1}\mathbf{L} = \mathbf{I} - \mathbf{L}'(\mathbf{L}\mathbf{L}' + \boldsymbol{\Psi})^{-1}\mathbf{L}$$

Cont.

- Então, pode-se usar como preditor para F_i :

$$\hat{F}_i^* = L' (LL' + \Psi)^{-1} (X_i - \mu)$$

- Substituindo-se Ψ , L e μ pelos respectivos estimadores, vem que:

$$\hat{F}_i = \hat{L}' \left(\hat{L}\hat{L}' + \hat{\Psi} \right)^{-1} (X_i - \bar{X})$$

- Note ainda que $\mathcal{E}(\hat{F}_i^*) = L' (LL' + \Psi)^{-1} (\mathcal{E}(X_i) - \mu) = \mathbf{0}$.

Cont.

- Além disso,

$$\begin{aligned} \text{Cov}(\widehat{\mathbf{F}}_i^*) &= \mathbf{L}' (\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} \text{Cov}(\mathbf{X}_i - \boldsymbol{\mu}) (\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} \mathbf{L} \\ &= \mathbf{L}' (\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} (\mathbf{L}\mathbf{L}' + \mathbf{\Psi}) (\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} \mathbf{L} \\ &= \mathbf{L}' (\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} \mathbf{L} \\ &= \mathbf{L}' \left(\mathbf{\Psi}^{-1} - \mathbf{\Psi}^{-1} \mathbf{L} (\mathbf{I} + \mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L})^{-1} \mathbf{L}' \mathbf{\Psi}^{-1} \right) \mathbf{L} \\ &= \mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L} - \mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L} (\mathbf{I} + \mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L})^{-1} \mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L} \end{aligned}$$

- Se \mathbf{L} e $\mathbf{\Psi}$ tiverem sido estimados por mv, então $\mathbf{L}' \mathbf{\Psi}^{-1} \mathbf{L} = \mathbf{\Delta}$, logo:

$$\text{Cov}(\widehat{\mathbf{F}}_i^*) = \mathbf{\Delta} - \mathbf{\Delta} (\mathbf{I} + \mathbf{\Delta})^{-1} \mathbf{\Delta}$$

Cont.

- Portanto, $\mathcal{E}(\hat{\mathbf{F}}_i) \approx \mathbf{0}$ e $\text{Cov}(\hat{\mathbf{F}}_i) \approx \hat{\mathbf{L}}' \left(\hat{\mathbf{L}}\hat{\mathbf{L}}' + \hat{\mathbf{\Psi}} \right) \hat{\mathbf{L}}$.
- No caso de \mathbf{L} e $\mathbf{\Psi}$ terem sido estimados por mv, então $\text{Cov}(\hat{\mathbf{F}}_i) \approx \hat{\mathbf{\Delta}} - \hat{\mathbf{\Delta}} \left(\mathbf{I} + \hat{\mathbf{\Delta}} \right)^{-1} \hat{\mathbf{\Delta}}$.
- Para ambos os métodos, quando se se considera as variáveis padronizadas, além de se substituir $\mathbf{\Sigma}$ por ρ , utiliza-se $\mathbf{z}_i = \mathbf{D}^{-1/2} (\mathbf{x}_i - \bar{\mathbf{x}})$ em que \mathbf{D} é uma matriz diagonal com as variâncias amostrais de \mathbf{X}_i em sua diagonal.
- As estimativas (valores preditos) são obtivos substituindo-se os estimadores pelas respectivas estimativas.

MQP \times MR

- Utilizando-se a relação $\mathbf{L}'(\mathbf{L}\mathbf{L}' + \mathbf{\Psi})^{-1} = (\mathbf{I} + \mathbf{L}'\mathbf{\Psi}^{-1}\mathbf{L})^{-1}\mathbf{L}'\mathbf{\Psi}^{-1}$, pode-se provar que

$$\widehat{\mathbf{F}}_i^{*MQP} = (\mathbf{L}'\mathbf{\Psi}^{-1}\mathbf{L})^{-1} (\mathbf{I} + \mathbf{L}'\mathbf{\Psi}^{-1}\mathbf{L}) \widehat{\mathbf{F}}_i^{*R} = \left(\mathbf{I} + (\mathbf{L}'\mathbf{\Psi}^{-1}\mathbf{L})^{-1} \right) \widehat{\mathbf{F}}_i^{*R}$$

- Sob as estimativas de mv de \mathbf{L} e $\mathbf{\Psi}$, temos que

$$\widehat{\mathbf{F}}_i^{*MQP} = (\mathbf{I} + \mathbf{\Delta}^{-1}) \widehat{\mathbf{F}}_i^{*R}$$

Resumo: estimação de L e Ψ

- A partir de uma matriz de dados $\mathbf{X}_{(n \times p)}$, calcula-se a matriz \mathbf{S}^2 (covariâncias) ou a matriz \mathbf{R} (correlações).
- Componentes principais: calcula-se $(\tilde{\lambda}_1, \tilde{\mathbf{e}}_1)'$, ..., $(\tilde{\lambda}_m, \tilde{\mathbf{e}}_m)'$ através de \mathbf{S}^2 ou \mathbf{R} . Preferencialmente \mathbf{R} .
- Teremos $\tilde{\mathbf{L}}$ e $\tilde{\Psi}$ a partir de \mathbf{S}^2 e $\tilde{\mathbf{L}}_Z$ e $\tilde{\Psi}_Z$ a partir de \mathbf{R} . Em

$$\text{que } \begin{bmatrix} \tilde{\psi}_1 & 0 & \dots & 0 \\ 0 & \tilde{\psi}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \tilde{\psi}_n \end{bmatrix} = \tilde{\Psi} = \text{diag}(\mathbf{S}^2 - \tilde{\mathbf{L}}\tilde{\mathbf{L}}'), \text{ ou}$$
$$\tilde{\Psi}_Z = \text{diag}(\mathbf{R} - \tilde{\mathbf{L}}_Z\tilde{\mathbf{L}}_Z')$$

Resumo

- A partir de uma matriz de dados $\mathbf{X}_{(n \times p)}$, maximiza-se à verossimilhança em função de $\boldsymbol{\mu}$, \mathbf{L} , $\boldsymbol{\Psi}$, com as devidas restrições de identificabilidade. Obtem-se assim, $\tilde{\boldsymbol{\mu}} = \bar{\mathbf{x}}$, $\tilde{\mathbf{L}}$, $\tilde{\boldsymbol{\Psi}}$ e $\tilde{\boldsymbol{\Sigma}} = \tilde{\mathbf{L}}\tilde{\mathbf{L}}' + \tilde{\boldsymbol{\Psi}}$.
- Para a matriz de correlações trabalhamos com \mathbf{R} , ou seja, utilizamos \mathbf{Z} , obtendo-se: $\tilde{\mathbf{L}}_Z$, $\tilde{\boldsymbol{\Psi}}_Z$ e $\tilde{\mathbf{R}} = \tilde{\mathbf{L}}_Z\tilde{\mathbf{L}}_Z' + \tilde{\boldsymbol{\Psi}}_Z$, lembrando que $\mathbf{Z} = (\mathbf{X} - \bar{\mathbf{x}})\tilde{\mathbf{D}}^{-1/2}$, em que $\tilde{\mathbf{D}} = \text{diag}(s_1^2, \dots, s_p^2)$.

Resumo

- Considera-se $\tilde{\mathbf{L}}$, $\tilde{\Psi}$ ou $\tilde{\mathbf{L}}_Z$, $\tilde{\Psi}_Z$, como valores conhecidos (obtidas via método das componentes principais ou máxima verossimilhança).

- MQP:

- $\tilde{\mathbf{F}}_i = \left(\tilde{\mathbf{L}}' \tilde{\Psi}^{-1} \tilde{\mathbf{L}} \right)^{-1} \tilde{\mathbf{L}}' \tilde{\Psi}^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$

- $\tilde{\mathbf{F}}_i = \left(\tilde{\mathbf{L}}_Z' \tilde{\Psi}_Z^{-1} \tilde{\mathbf{L}}_Z \right)^{-1} \tilde{\mathbf{L}}_Z' \tilde{\Psi}_Z^{-1} \mathbf{z}_i$

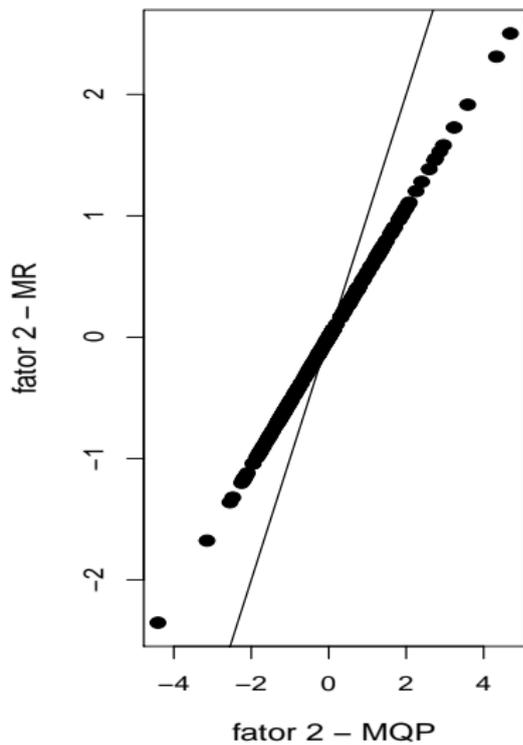
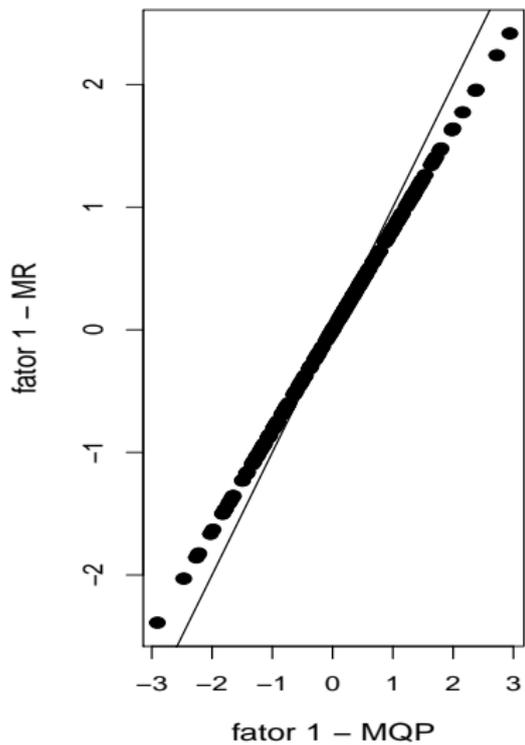
- Regressão:

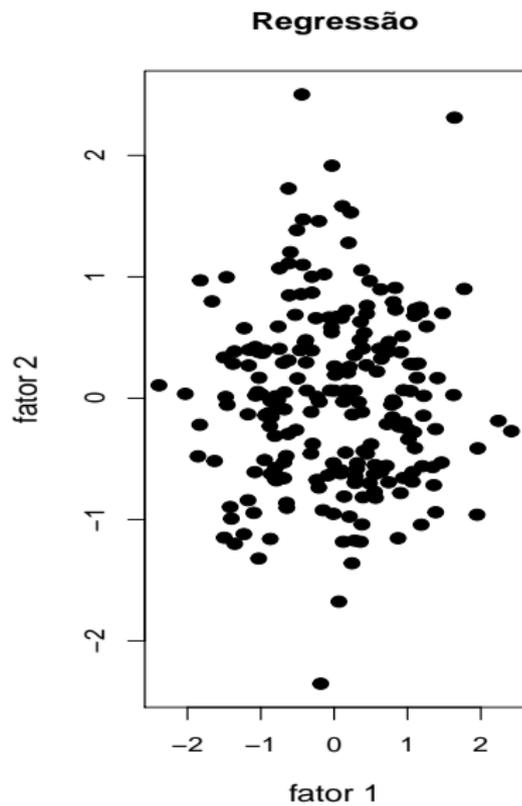
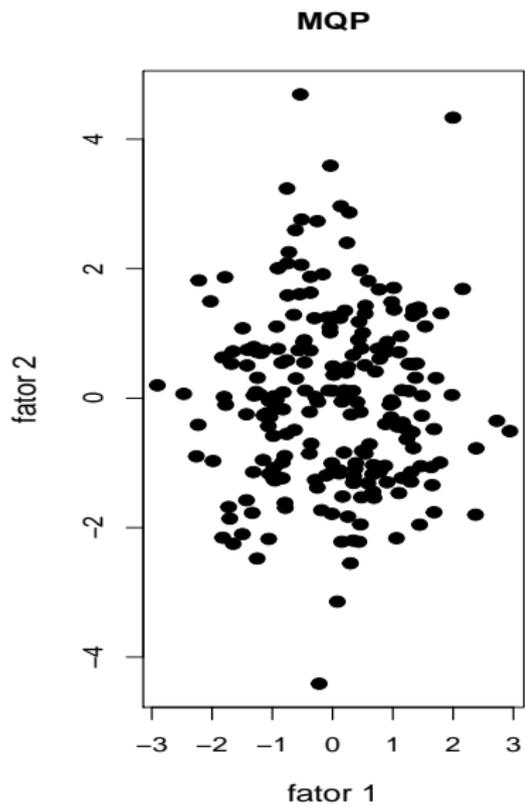
- $\tilde{\mathbf{F}}_i = \tilde{\mathbf{L}}' \left(\tilde{\mathbf{L}} \tilde{\mathbf{L}}' + \tilde{\Psi} \right)^{-1} (\mathbf{x}_i - \bar{\mathbf{x}})$.

- $\tilde{\mathbf{F}}_i = \tilde{\mathbf{L}}_Z' \left(\tilde{\mathbf{L}}_Z \tilde{\mathbf{L}}_Z' + \tilde{\Psi}_Z \right)^{-1} \mathbf{z}_i$.

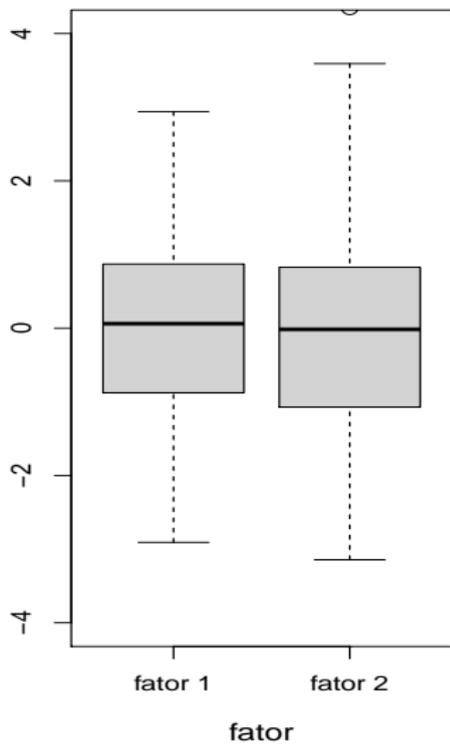
Voltando ao Exemplo 5

Estatística	MQP		Regressão	
	Fator 1	Fator 2	Fator 1	Fator 2
Média	0,00	0,00	0,00	0,00
Var.	1,18	1,64	0,85	0,61
DP	1,09	1,28	0,92	0,78
Mínimo	-3,56	-3,94	-3,01	-2,41
Mediana	0,03	0,03	0,02	0,02
Máximo	2,79	3,02	2,37	1,84

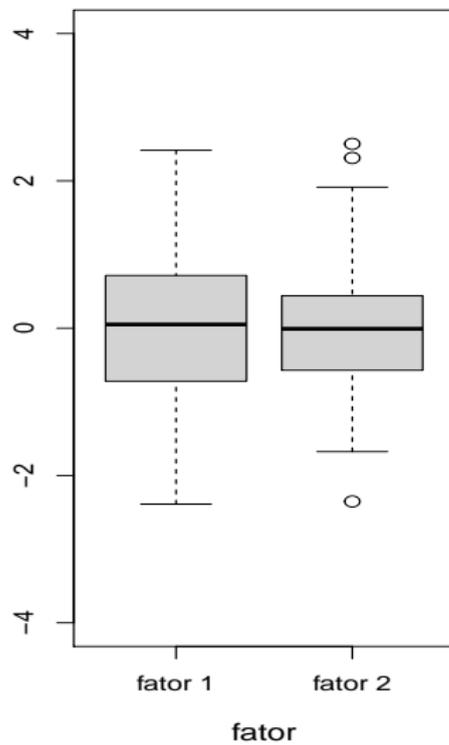


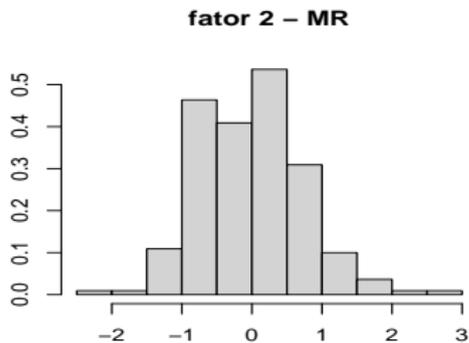
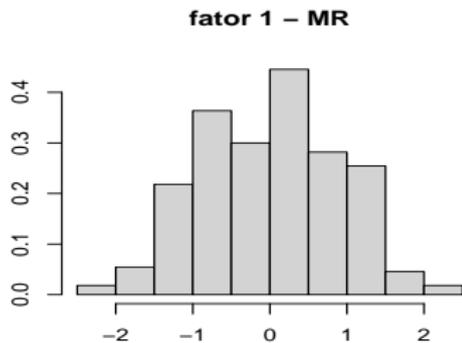
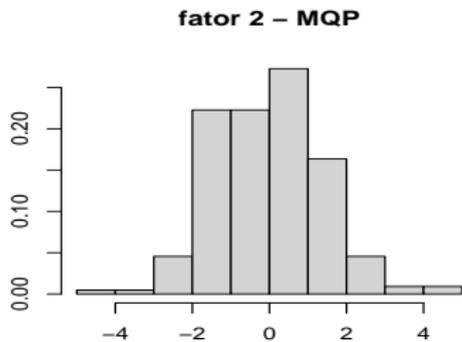
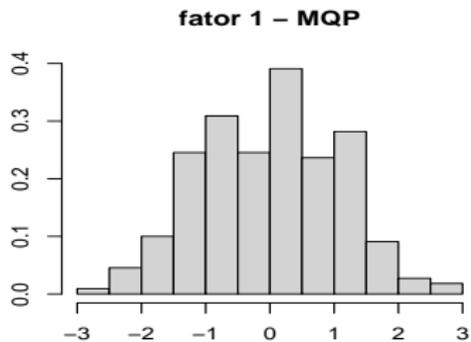


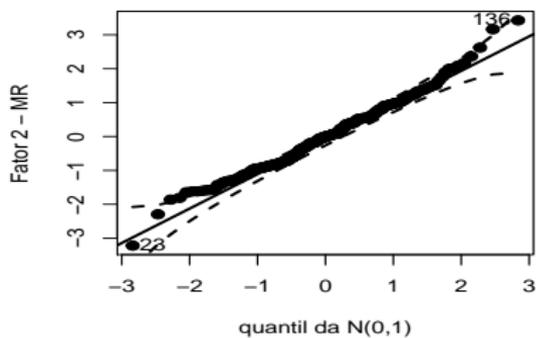
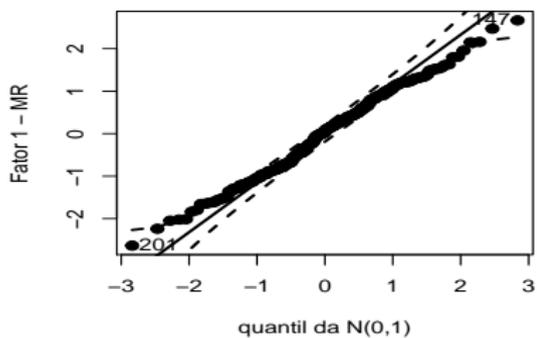
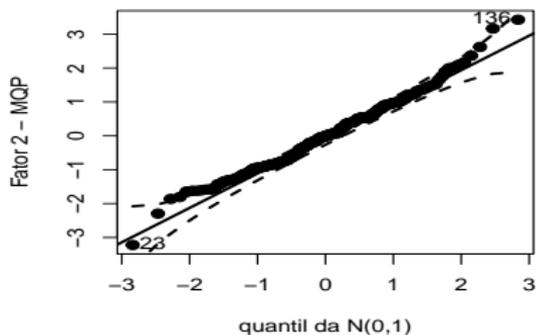
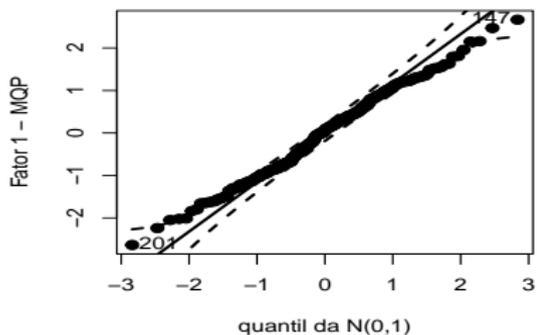
MQP

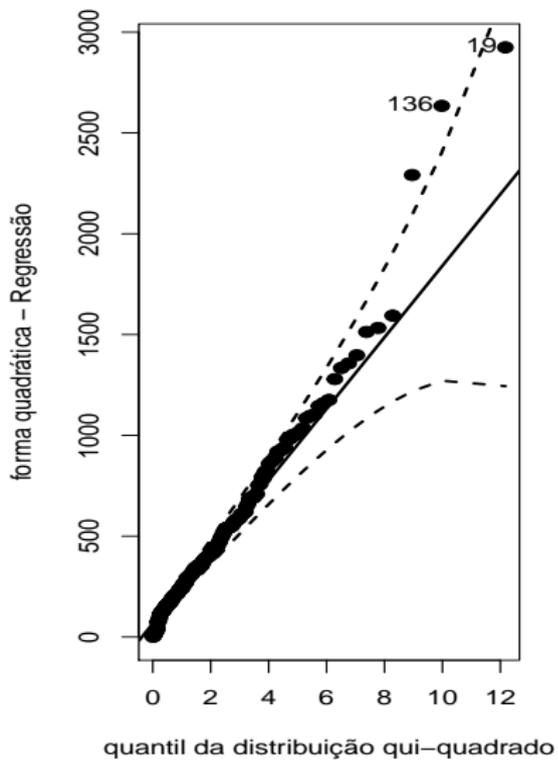
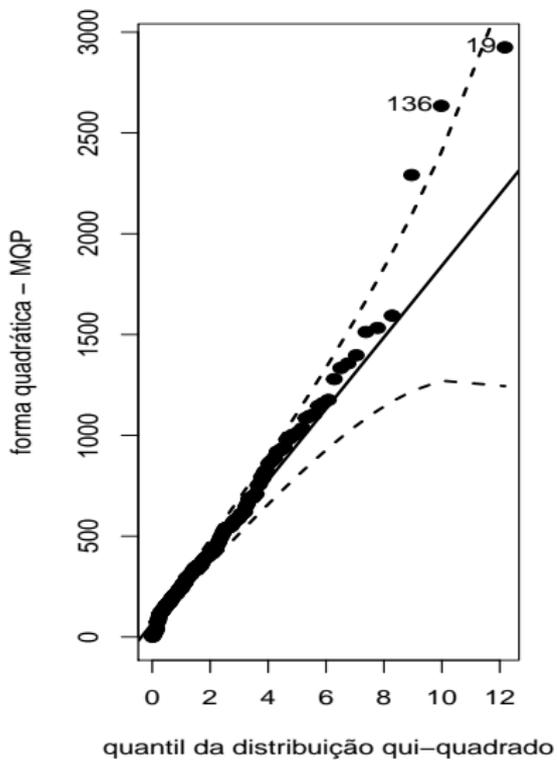


Regressão









Rotação dos fatores (cargas fatoriais)

- Vimos anteriormente que se $T_{(m \times m)}$ é uma matriz ortogonal ($T^{-1} = T'$) e definindo $L^* = LT$, então, tem-se que:

$$L^*(L^*)' + \Psi = (LT)(LT)' + \Psi = LTT'L' + \Psi = LL' + \Psi$$

- Ou seja, para uma dada estimativa das cargas fatoriais, uma outra pode ser obtida de modo a preservar a representação originalmente obtida para a matriz de covariâncias (correlações).
- Naturalmente, se usarmos L^* ao invés de L , outros valores para os escores fatoriais serão obtidos.

Cont.

- Há também rotações não ortogonais, por exemplo as oblíquas, as quais, quase nunca, fornecem a mesma representação da matriz de covariâncias (correlações).
- De **álgebra matricial**, sabemos que uma transformação (ortogonal ou não) corresponde a uma rotação (ou reflexão) dos eixos coordenados.
- Por esta razão, uma transformação (ortogonal ou não) das cargas dos fatores, e a transformação subsequente dos fatores, é chamada de rotação dos fatores.

Cont.

- Como as cargas originais dos fatores podem não ser (facilmente) interpretáveis, uma prática comum é rotacioná-las até que uma estrutura de cargas “mais simples” seja atingida.
- Uma situação ideal é buscar por um padrão de cargas tal que cada variável tenha cargas altas em um único fator (ou grupo de fatores) e cargas pequenas ou moderadas nos demais fatores.
- Porém nem sempre é possível chegar à essa estrutura simples.

Cont.

- Quando $m=2$ ou os fatores comuns são considerados dois de cada vez, a transformação para uma estrutura mais simples pode frequentemente ser determinada graficamente.
- Os fatores comuns (não correlacionados) são olhados como vetores unitários no plano Cartesiano cujos eixos coordenados correspondem a cada um dos dois fatores.
- Um gráfico dos pares das cargas dos dois fatores $(\tilde{l}_{i1}, \tilde{l}_{i2}), i = 1, 2, \dots, p$ produz p pontos, cada um correspondendo à uma variável.
- Os eixos coordenados podem, então, ser “visualmente rotacionados” através um ângulo, digamos ϕ .

Exemplo com dois fatores: rotação ortogonal

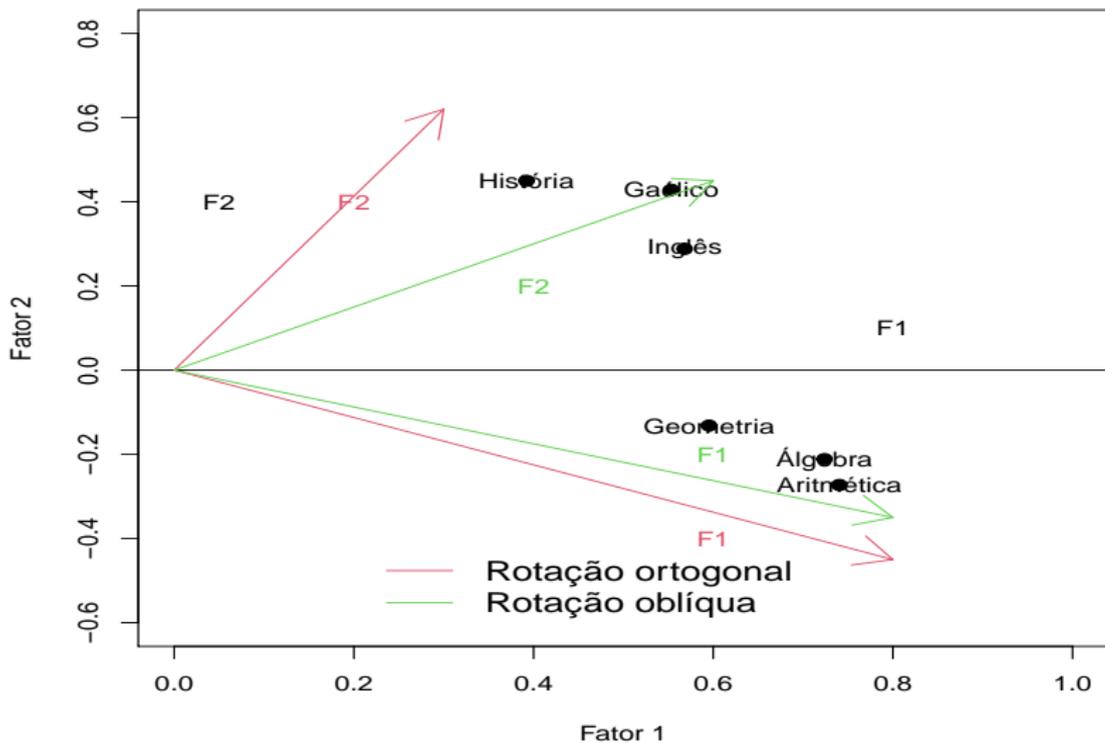
- Suponha que, para um dado exemplo, decidimos considerar dois fatores. Podemos definir a seguinte transformação ortogonal

$$\widehat{\mathbf{L}}_{(p \times 2)}^* = \widehat{\mathbf{L}}_{(p \times 2)} \mathbf{T}_{(2 \times 2)}$$

$$\text{em que } \left\{ \begin{array}{l} \mathbf{T} = \begin{bmatrix} \cos\phi & \text{sen}\phi \\ -\text{sen}\phi & \cos\phi \end{bmatrix}, \text{ rotação no sentido horário} \\ \mathbf{T} = \begin{bmatrix} \cos\phi & -\text{sen}\phi \\ \text{sen}\phi & \cos\phi \end{bmatrix}, \text{ rotação no sentido anti-horário} \end{array} \right.$$

- Estimativas das cargas fatoriais (Exemplo 5) : máxima verossimilhança

Var.	Fator 1	Fator 2	Comun. ($\sum_{j=1}^m l_{ij}^2$)	Var. Específ. (ψ_i)
Gaélico	0,553	0,429	0,490	0,510
Inglês	0,568	0,288	0,406	0,594
História	0,392	0,450	0,356	0,644
Aritmética	0,740	-0,273	0,623	0,377
Álgebra	0,724	-0,211	0,569	0,431
Geometria	0,595	-0,132	0,372	0,628



- Estimativas das cargas fatoriais, rotação ortogonal (Exemplo 5) :
máxima verossimilhança

Var.	Fator 1	Fator 2	Comun. ($\sum_{j=1}^m l_{ij}^2$)	Var. Específ. (ψ_i)
Gaélico	0,366	0,701	0,625	0,375
Inglês	0,568	0,316	0,422	0,578
História	0,165	0,581	0,365	0,635
Aritmética	0,815	-0,020	0,665	0,335
Álgebra	0,772	0,049	0,599	0,401
Geometria	0,617	0,072	0,386	0,614

- Fator 1: fator relativo à Matemática. Fator 2: fator relativo às ciências humanas. Rotação considerando $\phi = 20^\circ$ na matriz T anterior (rotação no sentido horário).

Exemplo com dois fatores

- Nesse caso ($m=2$), conglomerados de variáveis são freqüentemente aparentes graficamente, e esses conglomerados permitem-nos identificar os fatores comuns, sem ter que inspecionar as magnitudes das cargas.
- Por outro lado, quando $m \geq 2$, padrões não são facilmente visualizados e as magnitudes das cargas rotacionadas devem ser verificadas para se encontrar uma interpretação útil dos dados originais.
- A escolha de uma matriz ortogonal que satisfaz uma “medida analítica de estrutura” simples será considerada.

Rotação dos fatores (cargas fatoriais)

- Rotação ortogonal: Seja $T_{(m \times m)}$ uma matriz ortogonal ($T^{-1} = T'$) e defina $L^* = LT$, assim:

$$L^*(L^*)' + \Psi = (LT)(LT)' + \Psi = LTT'L' + \Psi = LL' + \Psi$$

- Rotação oblíqua: Seja $T_{(m \times m)}$ uma matriz não-ortogonal e defina $L^* = LT$, assim:

$$L^*(L^*)' + \Psi = (LT)(LT)' + \Psi = LTT'L' + \Psi \neq LL' + \Psi$$

Rotações ortogonais e oblíquas

- Rotação ortogonal:
 - Corresponde à uma rotação rígida (tende a manter a estrutura dos ângulos entre os vetores que representam os fatores).
 - Preserva a representação da matriz de covariâncias obtida pela solução original.
 - Preserva a ortogonalidade dos fatores.
- Rotações oblíquas:
 - Corresponde à uma rotação não-rígida (tende a não manter a estrutura dos ângulos entre os vetores que representam os fatores).
 - Não preserva a representação da matriz de covariâncias obtida pela solução original.
 - Não preserva a ortogonalidade dos fatores.

Rotação ortogonal varimax

- O método de rotação **varimax** usa como critério uma medida analítica (simples) denominada de critério varimax ou critério normal varimax.
- Sejam $\tilde{l}_{ij}^* = \tilde{l}_{ij}/\tilde{h}_i, i = 1, \dots, p; j = 1, \dots, m$ as cargas fatoriais a serem rotacionadas escalonadas pela comunalidade. O método varimax seleciona a matriz \mathbf{T} de modo a maximizar

$$V = \frac{1}{p} \sum_{j=1}^m \left[\sum_{i=1}^p \tilde{l}_{ij}^{*4} - \left(\sum_{i=1}^p \tilde{l}_{ij}^{*2} \right)^2 / p \right]$$

- Essencialmente $V \propto \sum_{j=1}^m (\text{variância do quadrado das cargas escalonadas do } j\text{-ésimo fator}).$

Cont.

- Escalonar os coeficientes rotacionados \tilde{l}_{ij}^* tem o efeito de atribuir a variáveis com comunalidades pequenas mais peso na determinação de uma estrutura simples.
- Após a transformação \mathbf{T} ter sido determinada, as cargas \tilde{l}_{ij}^* são multiplicadas por \tilde{h}_i para que as comunalidades originais sejam preservadas.
- Portanto, espera-se encontrar tanto grupos de coeficientes de elevada magnitude como grupos de coeficientes de magnitude desprezível em cada coluna da matriz de cargas rotacionadas.

- Estimativas das cargas fatoriais, rotação varimax (Exemplo 5) :
máxima verossimilhança

Var.	Fator 1	Fator 2	Comun. ($\sum_{j=1}^m l_{ij}^2$)	Var. Específ. (ψ_i)
Gaélico	0,250	0,750	0,625	0,375
Inglês	0,511	0,402	0,422	0,578
História	0,071	0,600	0,365	0,635
Aritmética	0,808	0,110	0,665	0,335
Álgebra	0,755	0,170	0,599	0,401
Geometria	0,598	0,169	0,386	0,614

- Fator 1: fator relativo à Matemática. Fator 2: fator relativo às ciências humanas.

Rotações oblíquas

- Rotações ortogonais são apropriadas para um modelo fatorial no qual os fatores são supostos não correlacionados.
- Em algumas áreas de pesquisa é comum (de interesse) considerar rotações oblíquas (não ortogonais), além das não ortogonais.
- As primeiras (oblíquas) são sugeridas após olhar as cargas estimadas e não a partir dos postulados básicos do modelo. Apesar disso, uma rotação oblíqua pode ser útil.

Rotações oblíquas

- Se olhamos os m fatores comuns como eixos coordenados, o ponto com as coordenadas das cargas da j -ésima variável sobre os m fatores comuns representa a posição dessa variável no espaço fatorial (como no gráfico anterior relativo ao exemplo de rotação).
- Supondo que as variáveis são agrupadas em conglomerados não sobrepostos, uma rotação ortogonal para uma estrutura mais simples corresponde a uma rotação rígida dos eixos coordenados tal que os eixos, após a rotação, passam tão próximos dos conglomerados quanto possível.

Rotações oblíquas

- Uma rotação oblíqua, para simplificar a estrutura das cargas, corresponde à uma rotação não rígida do sistema de coordenadas tal que os eixos rotacionados (não mais ortogonais dois a dois) passam (próximos) dos conglomerados.
- Uma rotação oblíqua busca expressar cada variável em termos de um número mínimo de fatores, preferivelmente um único fator.

Rotações oblíqua: promax

- Para rotações oblíquas, o método mais popular é o método promax que tem a vantagem de ser rápido e conceitualmente simples.
- O método busca ajustar uma matriz de cargas fatoriais (\mathbf{L}) (alvo) que tenha uma estrutura simples.
- A matriz alvo pode ser baseada, por exemplo, na matriz gerada pelo método de varimax.

- Estimativas das cargas fatoriais, rotação promax (Exemplo 5) :
máxima verossimilhança

Var.	Fator 1	Fator 2	Comun. ($\sum_{j=1}^m l_{ij}^2$)	Var. Específ. (ψ_i)
Gaélico	-0,004	0,793	0,629	0,371
Inglês	0,417	0,324	0,279	0,721
História	-0,147	0,669	0,469	0,531
Aritmética	0,859	-0,090	0,746	0,254
Álgebra	0,777	-0,007	0,604	0,396
Geometria	0,603	0,034	0,364	0,636

- Fator 1: fator relativo às Ciências Matemáticas. Fator 2: fator relativo à Ciências Humanas.

- Comparação entre as estimativas (originais e rotacionadas)
(Exemplo 5) : máxima verossimilhança

Var.	Original		Varimax		Promax	
	Fator 1	Fator 2	Fator 1	Fator 2	Fator 1	Fator 2
Gaélico	0,583	0,533	0,250	0,750	-0,004	0,793
Inglês	0,641	0,103	0,511	0,402	0,417	0,324
História	0,354	0,490	0,071	0,600	-0,147	0,669
Aritmética	0,760	-0,297	0,808	0,110	0,859	-0,090
Álgebra	0,742	-0,218	0,755	0,170	0,777	-0,007
Geometria	0,604	-0,143	0,598	0,169	0,603	0,034

Predição dos escores fatoriais

- De posse de alguma solução rotacionada, a predição dos fatores é feita de igual modo ao que foi anteriormente apresentado.
- Os fatores preditos, obtidos usando-se uma rotação ortogonal das cargas fatoriais, continuam sendo ortogonais.
- Os fatores preditos, obtidos usando-se uma rotação oblíqua das cargas fatoriais, já não mais são ortogonais.
- No gráfico a seguir temos os escores preditos a partir das cargas rotacionadas via varimax e promax.

