



UNIVERSIDADE FEDERAL DO PARÁ
INSTITUTO DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM MATEMÁTICA E ESTATÍSTICA

CURVAS DE CRESCIMENTO PARA
DADOS LONGITUDINAIS MULTIDIMENSIONAIS

Miguel Trócolis Lemos dos Santos

Belém
2017

Miguel Trócolis Lemos dos Santos

**Curvas de Crescimento para
Dados Longitudinais Multidimensionais**

Dissertação de Conclusão de Curso
apresentado ao Programa de Pós-
Graduação em Matemática e Es-
tatística da Universidade Federal do
Pará para a obtenção do Título de
Mestre em Estatística.

**Belém
2017**

Dados Internacionais de Catalogação - na - Publicação (CIP)
UFPA - Biblioteca Central

Santos, Miguel Trócolis Lemos,1991

Curvas de Crescimento para dados Longitudinais Multidimensionais/
Miguel Trócolis Lemos dos Santos. - 2017.

Orientador: Héilton Ribeiro Tavares

Dissertação (Mestrado) – Universidade Federal do Pará, Instituto de Ciências Exatas e Naturais, Programa de Pós-Graduação em Matemática e Estatística, Belém, 2017.

1. Estatística matemática. 2. Teoria da Resposta ao Item. 3. Curvas de crescimento. 4. Análise Multidimensional. I. Título

CDD 23 ed. 519.5

Miguel Trócolis Lemos dos Santos

**CURVAS DE CRESCIMENTO PARA DADOS LONGITUDINAIS
MULTIDIMENSIONAIS**

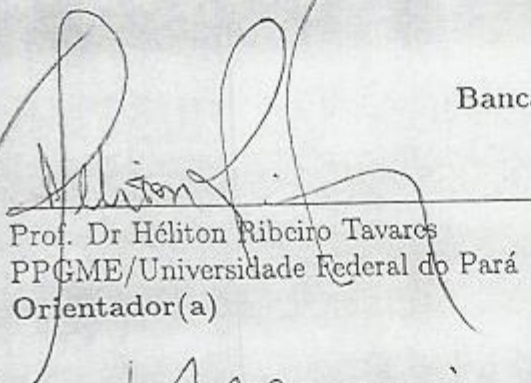
Esta Dissertação, foi julgada e aprovada, para a obtenção do grau de Mestre em Estatística, no Programa de Pós-Graduação em Matemática e Estatística da Universidade Federal do Pará.

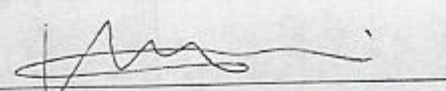
Belém, 29 de junho de 2017

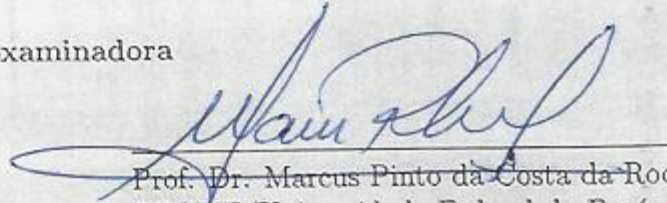
João Marcelo Brazão Protázio

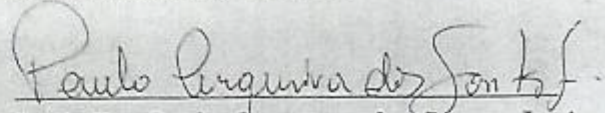
Prof. Dr. João Marcelo Brazão Protázio
(Coordenador(a) do Programa de Pós-Graduação em Matemática e Estatística - UFPA).

Banca Examinadora


Prof. Dr Héilton Ribeiro Tavares
PPGME/Universidade Federal do Pará
Orientador(a)


Prof. Dr. Valcir João da Cunha Farias
PPGME/PROFMAT/Universidade Federal do Pará
Examinador(a) Interno


Prof. Dr. Marcus Pinto da Costa da Rocha
PPGME/Universidade Federal do Pará
Examinador(a) Interno


Prof. Dr. Paulo Cerqueira dos Santos Junior
FAEST/Universidade Federal do Pará
Examinador(a) Externo

Resumo

Neste trabalho apresentamos um método para estimar mudanças nos vetores de médias das distribuições de traços latentes multivariados de grupos de examinados através de curvas de crescimento linear e não linear, no contexto da teoria da resposta ao item (IRT). Nossa abordagem permite trabalhar dentro de uma família de modelos logísticos de IRT e independe da etapa estimar os traços latentes individuais. São propostos quatro modelos para representar a verdadeira relação entre o vetor de médias da distribuição latente da população e uma variável explicativa, como o tempo ou o nível escolar. Diferentes estruturas de covariância são propostas para acomodar a dependência entre as habilidades sobre a variável explicativa. As equações de estimação são apresentadas para os parâmetros da curva de crescimento, bem como os parâmetros da estrutura de covariância das distribuições latentes multivariadas.

Palavras-chave: Teoria de Resposta ao Item, curvas de crescimento, estrutura de covariância, dados longitudinais, análise multidimensional.

Abstract

In this we present a method to estimate changes in the vector means of multivariate latent trait distributions of groups of examinees through linear and non-linear growth curves, within the framework of item response theory (IRT). Our approach allows working within a family of logistic models of IRT and it does not depend on estimating the individual latent traits. Four models are proposed to represent the true relationship between the population latent mean trait and an explanatory variable, such as time or grade. Different covariance structures are proposed to accommodate the dependence among the multivariate components and the abilities over the explanatory variable. The estimation equations are presented for the growth curve parameters, as well as the covariance parameters of the multivariate latent trait distributions.

Keywords: Item Response Theory, growth curve, longitudinal data, multidimensional analysis.

Sumário

1	Introdução	8
2	Modelo Probabilístico	10
2.1	Modelo Probabilístico da Teoria da Resposta ao Item	10
2.1.1	Parâmetro de discriminação do item	11
2.1.2	Parâmetro de dificuldade do item	11
2.1.3	Parâmetro c_i	12
2.1.4	Curva característica do Item	13
2.1.5	Distribuição de θ	13
2.1.6	Escala de Habilidade	13
2.1.7	Múltiplas Habilidades	14
2.1.8	Modelo Logístico	15
2.2	Modelo Probabilísticos: Análise Longitudinal	15
2.2.1	Introdução	15
2.2.2	Matriz de Correlação	16
2.2.3	Curvas de crescimento: Múltiplas habilidades	16
2.2.4	Definindo f	17
3	Processo de Estimação	19
3.1	Estimação por Máxima Verossimilhança	19
3.1.1	Introdução	19
3.1.2	Métodos da Máxima verossimilhança Marginal	20
3.2	Estimação das Habilidades	21
3.3	Raízes da log-verossimilhança	21
3.3.1	Método de Newton-Raphson	21
3.3.2	Método de Bock & Aitkin	23
4	Expressões para a estimação dos parâmetros	24
4.1	Parâmetros dos itens	24
4.1.1	1ª Derivada	24
4.1.2	2ª Derivada	27
4.2	Parâmetros Populacionais: Sem curvas Crescimento	30
4.2.1	1ª Derivada	30
4.2.2	2ª Derivada	32
4.3	Parâmetros Populacionais: Com curvas Crescimento	34
4.3.1	1ª Derivada	34
4.3.2	2ª Derivada	35
5	Considerações Finais	37

Capítulo 1

Introdução

Em uma grande variedade de problemas, há uma característica que direciona o comportamento de um objeto diante determinadas situações. Em muitos casos esta característica não pode ser medida diretamente, chamada assim de traço latente. A Teoria da Resposta ao Item (TRI) trata de modelos probabilísticos avançados que abrangem uma gama dessas situações afim de medir este traço latente, baseada nas respostas às situações impostas, chamadas itens. Por consequência, o estudo das propriedades e implicações dos itens também faz parte da TRI (ANDRADE, TAVARES & VALLE, 2000).

A TRI é usada amplamente, em todo mundo, no âmbito da avaliação educacional. No Brasil é usada no: Sistema Nacional de Avaliação da Educação Básica (SAEB), Exame Nacional do Ensino Médio (ENEM), Exame Nacional de Certificação de Competências para Jovens e Adultos (ENCCEJA), Programa Internacional de Avaliação de Alunos (PISA), Sistema Paraense de Avaliação Educacional (SisPAE), entre outros. A massificação do uso da TRI causou maior atração da comunidade acadêmica, cada vez mais disposta a estudá-la, desenvolvê-la e apontar aplicações diferenciadas.

A TRI vem sendo efetivada desde década de 50 com Lord (1952) que estipulou um modelo com respostas dicotômicas. O modelo de Lord possibilita a comparação entre as habilidades de indivíduos, mesmo com testes diferentes, além de definir parâmetros que caracterizam um item de forma bem coerente. A princípio, Lord trabalha apenas com dois parâmetros de item, porém, posteriormente ele acrescenta mais um parâmetro que traduz a possibilidade de um indivíduo com baixa habilidade (habilidade tendendo a menos infinito) acertar o item. Como função base para a probabilidade de um indivíduo acertar um determinado item, dado sua habilidade, Lord usou a função ogiva normal, porém Birnbaum (1968) a substituiu pela função logístico, algebricamente mais atraente.

Bock & Zimowski (1997) contribuíram com um modelo que analisa situações em que há grupos de indivíduos oriundos de populações diferentes. Cada população apresenta parâmetros distintos, que muitas vezes necessitam ser estimados. Agora pode-se comparar grupos e até mesmo itens dados a populações diferentes, tal processo é denominado equalização e exige condições específicas. (TAVARES, ANDRADE & VALLE, 2000).

Dado o modelo para trabalhar com várias populações, é possível imaginar outro

modelo que trata um mesmo grupo avaliado ao decorrer do tempo, ou seja, uma análise longitudinal. Pode-se até mesmo associar curvas de crescimento para a média das habilidades em relação ao tempo, possibilitando projeções. Para tanto, considera-se em cada instante que o grupo em questão é proveniente de uma população diferente, porém é óbvio que existe uma correlação entre as populações. Com esta visão pode-se citar Tavares (2001).

Reckase (2009) apresentou um modelo onde é considerado mais de uma habilidade essencial para responder um item. Por exemplo, no campo da educação, pode-se considerar para responder uma questão de matemática tanto a habilidade propriamente em matemática como uma outra habilidade relativa a interpretação de texto. O interesse por esta perspectiva é crescente na avaliação educacional.

Logo, a proposta deste trabalho é a elaboração teórica de um modelo probabilístico para aplicação de testes longitudinais com curvas de crescimento, para um único grupo, conforme apresentado por Tavares, entretanto, o diferencial será o uso de itens que exigem múltiplas habilidades. Também serão desenvolvidas as principais expressões necessárias para a aplicação do Método de Newton-Raphson.

Dado a proposta deste trabalho, seguem mais quatro capítulos. Abaixo está a descrição de cada um.

O Capítulo 2 tem como objetivo apresentar o modelo probabilístico. O modelo será gradativamente montado e alterado, tendendo para os objetivos do trabalho. Primeiramente será exposto, na Seção 2.1, o modelo da TRI para itens dicotômicos em múltiplas habilidades. Já na Seção 2.2, serão expostos os modelos com análise longitudinal e curva de crescimento para as médias populacionais.

O Capítulo 3 tem como objetivo expor as técnicas empregadas para a estimação dos parâmetros. Segue neste capítulo a função de log-verossimilhança marginal. Também seguem as expressões e técnicas suficientes para a estimação das habilidades dados os parâmetros populacionais e dos itens. Contudo, as expressões para a estimação dos parâmetros populacionais e dos itens, através do Método de Newton-Raphson serão deixadas para o próximo capítulo.

O Capítulo 4 tem como objetivo descrever a obtenção das principais expressões necessárias para a aplicação do algoritmo de Newton-Raphson. Ou seja, este capítulo trata da primeira e segunda derivada da função de log-verossimilhança marginal em relação aos parâmetros dos itens e populacionais.

No Capítulo 5 serão apresentadas as considerações finais sobre o trabalho e sugestões para trabalhos futuros.

Capítulo 2

Modelo Probabilístico

2.1 Modelo Probabilístico da Teoria da Resposta ao

Item

Dada a aplicação de um teste com I itens para um grupo de J indivíduos, será chamado U_{ji} a resposta pertinente ao indivíduo j para o item i . Para cada indivíduo j está associado o vetor de respostas $\mathbf{U}_j = (U_{j1}, \dots, U_{jI})'$. Já para todos os indivíduos tem-se o vetor $\mathbf{U}_{..} = (\mathbf{U}_1, \dots, \mathbf{U}_J)'$ que traduz toda a informação à priori.

Neste contexto, é de interesse estimar uma característica latente do indivíduo j , denotada por θ_j , de tal forma que a resposta a um item depende dessa característica, chamada habilidade. Por fim, assume-se o objetivo mais geral de estimar o vetor $\boldsymbol{\theta} = (\theta_1, \dots, \theta_J)'$, que traduz as habilidades de todos os J indivíduos alvos da experiência. Sobre $\boldsymbol{\theta}$, como exigência do modelo da TRI, é necessário que este vetor possa ser entendido como realizações de uma variável aleatória θ . Sendo assim, segue a notação para a função de densidade e distribuição respectivamente de tal variável: $g^{(0)}(\theta|\boldsymbol{\eta}^{(0)})$ e $G^{(0)}(\theta|\boldsymbol{\eta}^{(0)})$, tal que $\boldsymbol{\eta}^{(0)}$ é o vetor de parâmetros associado à distribuição.

O primeiro pressuposto para a construção do modelo é a independência entre as repostas de indivíduos diferentes, ou seja, o vetor aleatório \mathbf{U}_{j_1} é independente de \mathbf{U}_{j_2} , desde que $j_1 \neq j_2$.

O segundo pressuposto é a **independência condicional**, ou seja, as respostas a dois itens por um mesmo indivíduo são independentes, desde que a habilidade deste indivíduo seja dada. Matematicamente, $U_{ji_1}|\theta$ é independente de $U_{ji_2}|\theta$, desde que $i_1 \neq i_2$.

Para este trabalho é considerado uma resposta dicotômica, ou seja, incorreta, $U_{ji} = 0$; ou correta, $U_{ji} = 1$, e apenas um dos dois casos. O modelo probabilístico obviamente é

pautado na distribuição *Bernoulli*, assim $U_{ji} \sim \text{Bernoulli}(P_i)$, tal que:

$$P(U_{ji} = u_{ji} | \theta = \theta_j) = P_i^{u_{ji}} Q_i^{1-u_{ji}}, \quad (2.1)$$

sendo $Q_i = 1 - P_i$.

Agora a problemática resume-se em encontrar um valor apropriado para P_i . Como dito antes, a resposta depende da habilidade do indivíduo, assim P_i é função de θ . Para este modelo, as suposições feitas para P_i são as seguintes:

- P_i em função de θ é crescente, assim quanto maior a habilidade do indivíduo maior a probabilidade de acertar o item.
- Quando θ tende ao infinito P_i tende a um, assim um indivíduo com alta habilidade tem grande probabilidade de acertar o item.
- Quando θ tende ao menos infinito P_i tende a uma constante, denominada c_i , entendida como a probabilidade de um indivíduo com habilidade muito baixa acertar o item, representando o acerto ao acaso.

Uma função que atende às condições impostas é uma função de distribuição acumulada $F(\theta)$. Por agora, F não será definida. Entenda o modelo, por enquanto, como:

$$P_i = F(\theta),$$

tal que, F é uma função de probabilidade acumulada. Nas próximas três subseções, serão acrescentados os parâmetros dos itens ao modelo afim de diferenciar um item de outro. A ordem que os parâmetros serão adicionados aqui é irrelevante, pois a lógica que os torna necessários não prevê ordenação.

2.1.1 Parâmetro de discriminação do item

Note que, no modelo atual ($P_i = F(\theta)$), para uma variação de habilidade de um θ_1 para um θ_2 há sempre uma mesma variação de probabilidade para todos os itens (Veja Figura 2.1). Esse fato não traduz a maioria das situações reais.

Em muitos casos há itens que contemplam uma mudança significativa na probabilidade de acerto a uma determinada variação de habilidade e, com mesma variação de habilidade, outros nem tanto, sendo assim os últimos mais discriminatórios. Note que esta característica está ligada à taxa de variação de P_i , afim de alterá-la de acordo com o item, faz-se uma compressão (ou expansão) horizontal. Assim, altera-se o modelo:

$$P_i = F(a_i \theta),$$

em que a_i é chamado de parâmetro de discriminação do item i .

2.1.2 Parâmetro de dificuldade do item

Fixando uma habilidade para um indivíduo, no modelo atual ($P_i = F(a_i \theta)$), este indivíduo terá a mesma probabilidade de acertar qualquer item. Mais uma vez este fato não traduz a maioria das situações.

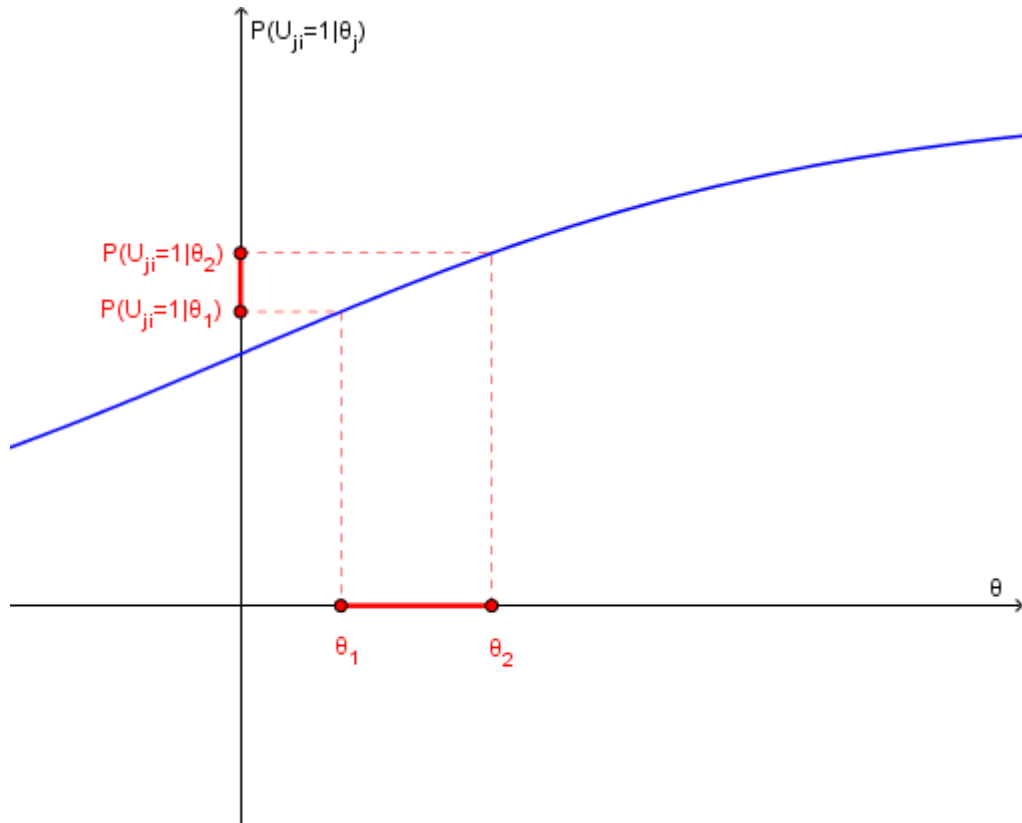


Figura 2.1: Variação da habilidade de θ_1 a θ_2 na função P

Em outras palavras, existem itens que exigem maior ou menor habilidade que outros para uma determinada probabilidade de acerto, assim o objetivo do ajuste agora está em diferenciar os itens por suas dificuldades. Uma solução imediata é deslocar toda a função horizontalmente b_i unidades de acordo com a dificuldade do item. Assim:

$$P_i = F(a_i(\theta - b_i)),$$

onde b_i é chamado parâmetro de dificuldade do item.

2.1.3 Parâmetro c_i

Vê-se no modelo dado até aqui o caso particular onde $c_i = 0$, o objetivo agora é ajustar o modelo para um c_i geral. Para subir a assíntota inferior, para um valor diferente de zero, é necessário somar tal valor a nossa função, o que gera um assíntota superior maior que um, afim de enquadrar a imagem da função no intervalo de c_i a 1 deve-se primeiramente multiplicar a função por $(1 - c_i)$, ou seja:

$$P_i = c_i + (1 - c_i)F(a_1(\theta - b_i)),$$

finalizando o modelo em relação aos parâmetros dos itens.

2.1.4 Curva característica do Item

A Curva Característica do Item (CCI) é obtida pela representação gráfica da probabilidade de acertar o item em função da habilidade. Segue um exemplar de CCI:

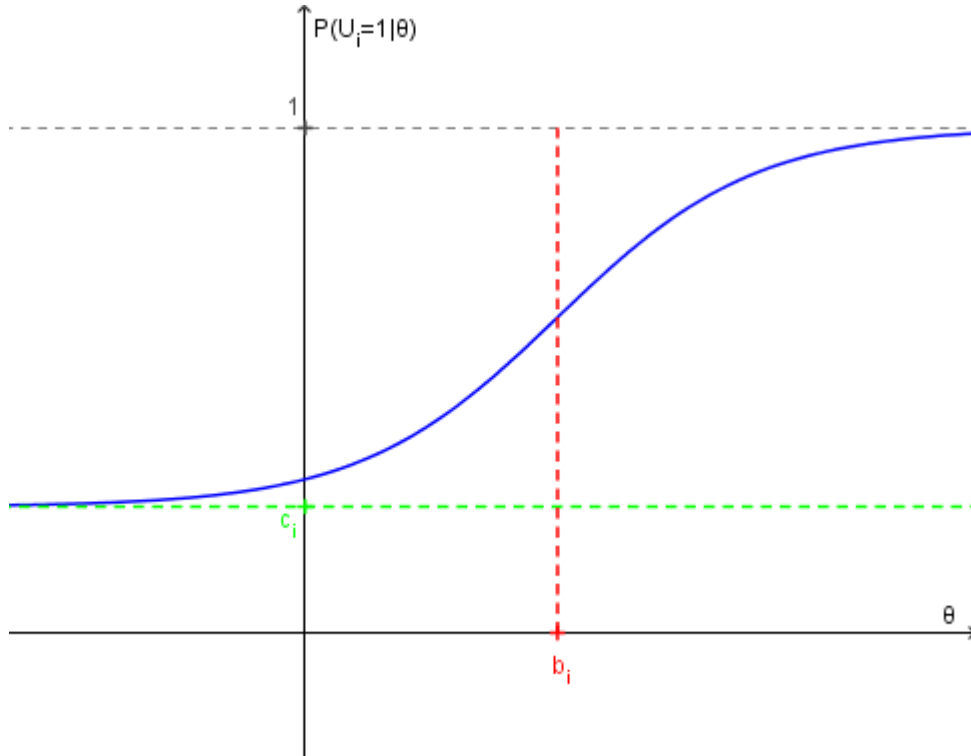


Figura 2.2: Curva Característica do Item

2.1.5 Distribuição de θ

Prosseguindo, resta definir uma distribuição conveniente para θ . É importante notar que a distribuição irá depender dos objetivos do pesquisador. Contudo, a distribuição Normal vem apresentando resultados satisfatórios em um setor abrangente de casos. Logo, $\theta \sim N(\mu, \sigma^2)$.

2.1.6 Escala de Habilidade

Chama-se escala o par ordenado (μ, σ) , em que μ é a média da população e σ o desvio padrão. Note que a escala depende tanto dos valores das habilidades, como dos parâmetros dos itens, devido ao processo de estimação. Ou seja, tem-se como resultados para estimação dos parâmetros e habilidades um resultado específico para uma determinada escala.

Contudo, uma habilidade, independentemente da escala, pode ser declarada como a quantidade x de desvios padrões que difere da média, Para obter o valor x de imediato na estimação basta considerar a escala $(0, 1)$.

No caso de uma escala não estabelecida, a escala (0,1) é suficiente. Isto acontece devido alguns destes traços latentes pesquisados não apresentarem métrica padrão, assim o pesquisador é livre pra estabelecer como padrão uma métrica que atribui a escala (0,1) para uma de suas populações. Mesmo com uma métrica já existente, é comum realizar os cálculos na escala (0,1) para apenas após o processo de estimação realizar uma conversão. O processo de conversão torna-se imediato:

$$\theta^* = \sigma\theta + \mu,$$

em que: θ^* é a habilidade na nova escala (μ, σ) , θ a habilidade na escala (0,1).

Outro caso é quando os parâmetros dos itens já estão estimados, ou seja, já há uma escala, portanto as habilidades estimadas estarão na mesma escala. O oposto também pode ocorrer, as habilidades são dadas e o interesse recai nos parâmetros dos itens, da mesma forma já há uma escala determinada.

2.1.7 Múltiplas Habilidades

Um caso mais abrangente da TRI é o caso multidimensional, em que a resposta a um item depende não só de uma habilidade, mas de K habilidades ou dimensões. Logo, nota-se $\theta_{kj}^{(1)}$ a habilidade de número k do indivíduo j , sendo assim o conjunto de habilidades, do tipo k , de todos os indivíduos é notado como $\boldsymbol{\theta}_k^{(1)} = (\theta_{k1}, \dots, \theta_{kJ})'$ que por sua vez será compreendido como uma amostra de uma variável aleatória $\theta_k^{(1)} \sim N(\mu_k^{(1)}, \sigma_k^{(1)2})$. Logo, $\boldsymbol{\theta}_{..}^{(1)} = (\boldsymbol{\theta}_1^{(1)}, \dots, \boldsymbol{\theta}_K^{(1)})'$ é uma amostra de uma normal K-variada.

A distribuição normal multivariada, acima, é definida pelos parâmetros $(\boldsymbol{\mu}^{(1)}, \boldsymbol{\Sigma}^{(1)})$, tal que defini-se $\sigma_{k_1 k_2}^{(1)}$ a covariância entre as habilidades $\theta_{k_1}^{(1)}$ e $\theta_{k_2}^{(1)}$, concluindo:

$$\boldsymbol{\mu}^{(1)} = (\mu_1^{(1)}, \dots, \mu_K^{(1)})',$$

$$\boldsymbol{\Sigma}^{(1)} = \begin{bmatrix} \sigma_1^{(1)2} & \sigma_{12}^{(1)} & \dots & \sigma_{1K}^{(1)} \\ \sigma_{21}^{(1)} & \sigma_2^{(1)2} & \dots & \sigma_{2K}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{K1}^{(1)} & \sigma_{K2}^{(1)} & \dots & \sigma_K^{(1)2} \end{bmatrix}.$$

Afim de obter a escala (0,1) para todas as habilidades $\theta_k^{(1)}$, faz-se $\mu_k^{(1)} = 0$ e $\sigma_k^{(1)2} = 1$ para todo k .

Um modelo conveniente para estrutura de covariância, afim de simplificar os cálculos, é o modelo uniforme, ou seja:

$$\boldsymbol{\Sigma}^{(1)} = \sigma^{(1)2} \begin{bmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_1 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_1 & \dots & 1 \end{bmatrix}, \quad (2.2)$$

tal que, $\sigma^{(1)2} = 1$, afim de respeitar a escala. Este modelo de fato é limitado, contudo é perfeito quando toma-se $K = 2$.

2.1.8 Modelo Logístico

Alguns modelos são apresentados para F ao longo da literatura. Frequentemente, pensa-se no modelo ogiva normal, que é dado por uma função de distribuição acumulada de uma variável normalmente distribuída com parâmetros $(0, 1)$. Contudo, o modelo normal não possui um resultado analítico para distribuição acumulada, assim Birnbaum (1968) apresentou como alternativa o modelo logístico como aproximação do normal. Portanto, no caso unidimensional:

$$P_i = c_i + (1 - c_i) \left\{ 1 + e^{-Da_i(\theta - b_i)} \right\}^{-1},$$

quando $D = 1,7$ tem-se uma aproximação ao modelo normal com parâmetros $(0, 1)$, ou seja, D é dado apenas como um artifício para aproximar as estimações dos parâmetros dos itens no modelo ogiva normal e logístico.

O modelo da expressão acima é denominado Modelo Logístico de três parâmetros (ML3). São casos particulares do ML3 o Modelo Logístico de dois parâmetros (ML2), onde $c_i = 0$, e o modelo logístico de um parâmetro (ML1), onde $c_i = 0$ e $a_i = 1$.

De modo análogo ao modelo unidimensional, obtêm-se o modelo logístico para o caso multidimensional:

$$P_i = c_i + (1 - c_i) \left\{ 1 + e^{-(a_{1i}\theta_1 + \dots + a_{Ki}\theta_K + d_i)} \right\}^{-1}. \quad (2.3)$$

Note que neste modelo desconsiderou-se o parâmetro D e foi dado d_i que correspondente a $-\sum_{k=1}^K a_{ki}b_{ki}$. Aproveitando, define-se $\zeta_i = (a_{1i}, \dots, a_{Ki}, d_i)'$, assim como $\zeta = (\zeta_1, \zeta_2, \dots, \zeta_I)'$.

2.2 Modelo Probabilísticos: Análise Longitudinal

2.2.1 Introdução

Por hora, desconsidere o caso multidimensional. Toma-se T teste aplicados cada um em momentos distintos a um mesmo grupo de indivíduos, dados por variações de tempo contínuas, tal que, para cada instante em que o teste foi aplicado é dado um valor t , onde $t = 1, \dots, T$. Portanto, será notado U_{jti} a resposta do indivíduo j no instante t ao item i , assim como define-se I_t o conjunto com os índices referentes aos itens aplicados no teste do instante t , T_i o conjunto de instantes t que foram aplicados o item i e por fim: $\mathbf{U}_{jt} = (U_{jti})'$, tal que, $i \in I_t$; $\mathbf{U}_{j.} = (\mathbf{U}_{j1}, \mathbf{U}_{j2}, \dots, \mathbf{U}_{jT})'$. Note que, muitas vezes espera-se a variação da habilidade desta população em função do tempo, ou seja, os parâmetros populacionais devem alterar-se com o valor de t . Sendo assim para cada t há uma média $(\mu_t^{(2)})$ e uma variância $(\sigma_t^{(2)})^2$. Será chamado $\theta_{jt}^{(2)}$ a habilidade de um indivíduo j no instante t , logo o vetor $\boldsymbol{\theta}_{.t}^{(2)} = (\theta_{1t}^{(2)}, \dots, \theta_{Jt}^{(2)})'$ será entendido como uma amostra de uma variável aleatória $\theta_t^{(2)}$.

Agora é pressuposta também uma independência temporal, ou seja, as respostas de um mesmo indivíduo em instantes diferentes são independentes, dado suas habilidades. Matematicamente, $U_{jt_1i_1}|\theta_{jt_1}$ é independente de $U_{jt_2i_2}|\theta_{jt_2}$, desde de que $t_1 \neq t_2$. É importante notar que no caso da aplicação na avaliação educacional, para admitir tal independência é necessário

$i_1 \neq i_2$ ou que o intervalo de tempo seja grande o suficiente para o aluno esquecer o item caso $i_1 = i_2$.

É comum nesta situação a modelagem através de uma curva que descreva o comportamento de $\theta_t^{(2)}$:

$$\theta_t^{(2)} = f(t|\boldsymbol{\alpha}^{(2)}) + \epsilon;$$

sendo ϵ entendido como um “erro aleatório”, este valor é concebido, em muitos casos, com distribuição normal $(0, \sigma_t^{(2)2})$, e f uma função em relação a t dado um vetor de parâmetros $\boldsymbol{\alpha}^{(2)} = (\alpha_1, \dots, \alpha_P)$.

Note que:

$$\begin{aligned} E(\theta_t^{(2)}) &= E(f(t|\boldsymbol{\alpha}^{(2)}) + \epsilon) \Rightarrow \\ \mu_t^{(2)} &= f(t|\boldsymbol{\alpha}^{(2)}). \end{aligned}$$

2.2.2 Matriz de Correlação

Sendo $\boldsymbol{\theta}^{(2)} = (\theta_1^{(2)}, \theta_2^{(2)}, \dots, \theta_T^{(2)})'$ um vetor aleatório, e dada as características de $\theta_t^{(2)}$ na subsecção anterior, $\boldsymbol{\theta}^{(2)}$ é uma variável aleatória com distribuição normal T-variada com parâmetros $\boldsymbol{\mu}^{(2)}$ e $\boldsymbol{\Sigma}^{(2)}$. Esta subsecção trata da estrutura para $\boldsymbol{\Sigma}^{(2)}$.

Existem muitas estruturas sugeridas para $\boldsymbol{\Sigma}^{(2)}$ porém a escolhida a qui é autorregressivo de ordem um, AR(1). A estrutura AR(1) considera uma mesma variância para todo θ_t , contudo, a correlação diminui exponencialmente em relação ao tempo, veja:

$$\boldsymbol{\Sigma}^{(2)} = \sigma^{(2)2} \begin{bmatrix} 1 & \rho_2 & \rho_2^2 & \dots & \rho_2^{T-1} \\ \rho_2 & 1 & \rho_2 & \dots & \rho_2^{T-2} \\ \rho_2^2 & \rho_2 & 1 & \dots & \rho_2^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_2^{T-1} & \rho_2^{T-2} & \rho_2^{T-3} & \dots & 1 \end{bmatrix}. \quad (2.4)$$

É necessário compor uma escala para uma das populações afim de servir como referência. Para tanto toma-se $\mu_1^{(2)} = 0$ e $\sigma^{(2)2}$.

2.2.3 Curvas de crescimento: Múltiplas habilidades

Partindo para o caso multidimensional, nota-se θ_{kjt} a habilidade k do indivíduo j no instante t , assim como $\boldsymbol{\theta}_{k.t} = (\theta_{k1t}, \dots, \theta_{kjt})'$, e ainda $\boldsymbol{\theta}_{.t} = (\boldsymbol{\theta}_{1.t}, \dots, \boldsymbol{\theta}_{K.t})$ que será uma amostra da variável aleatória $\boldsymbol{\theta}_t$, dado:

$$\boldsymbol{\theta}_t = f(t|\boldsymbol{\alpha}) + \boldsymbol{\epsilon}$$

A diferença é que agora $\boldsymbol{\epsilon} = (\epsilon_1, \dots, \epsilon_K)'$ tem distribuição normal K-variada. Portanto:

$$f(t|\boldsymbol{\alpha}) = \boldsymbol{\mu} = \begin{bmatrix} f_1(t|\boldsymbol{\alpha}_1) \\ \vdots \\ f_K(t|\boldsymbol{\alpha}_K) \end{bmatrix},$$

sendo que $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \dots, \boldsymbol{\alpha}_K)'$, ou seja, há um vetor $\boldsymbol{\alpha}_k = (\alpha_{1k}, \dots, \alpha_{Pk})'$ para cada habilidade k .

Note que foram compostas duas matrizes de covariância, para a análise multidimensional, matriz A , e outra para a análise longitudinal, matriz B . Veja:

$$A = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_1 \\ \rho_1 & 1 & \dots & \rho_1 \\ \vdots & \vdots & \ddots & \vdots \\ \rho_1 & \rho_1 & \dots & 1 \end{bmatrix}.$$

$$B = \begin{bmatrix} 1 & \rho_2 & \rho_2^2 & \dots & \rho_2^{T-1} \\ \rho_2 & 1 & \rho_2 & \dots & \rho_2^{T-2} \\ \rho_2^2 & \rho_2 & 1 & \dots & \rho_2^{T-3} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \rho_2^{T-1} & \rho_2^{T-2} & \rho_2^{T-3} & \dots & 1 \end{bmatrix}.$$

Por fim, será tomada uma variável aleatória $\boldsymbol{\theta} = (\theta_{11}, \dots, \theta_{1T}, \theta_{21}, \dots, \theta_{KT})'$ com distribuição normal TK -variada de parâmetros $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$, sendo:

$$\begin{aligned} \boldsymbol{\mu} &= (\mu_{11}, \dots, \mu_{KT})', \\ \boldsymbol{\Sigma} &= A \otimes B, \end{aligned}$$

assim o conjunto de todas as habilidades a ser estimado é uma amostra de $\boldsymbol{\theta}$.

Note que torna-se conveniente tomar $\boldsymbol{\Sigma}$ como o produto de kronecker, pois ao fixar um instante t obtém-se como matriz de covariância, para as habilidades, do tipo dada em (2.2); por outro lado tomando apenas um habilidade k obtém-se como matriz de covariância do tipo dado em (2.4). Ou seja, preserva-se os pontos tomados como base para o modelo.

Para finalizar a questão das notações defini-se: $\boldsymbol{\eta} = (\boldsymbol{\mu}', \rho_1, \rho_2)'$ ou $\boldsymbol{\eta} = (\boldsymbol{\alpha}', \rho_1, \rho_2)'$ no caso que se admita curvas de crescimento, $g(\boldsymbol{\theta}|\boldsymbol{\eta})$ e $G(\boldsymbol{\theta}|\boldsymbol{\eta})$ são funções de densidade e distribuição de $\boldsymbol{\theta}$ respectivamente, $\boldsymbol{\theta}_{.j}$ é o vetor com todas as habilidades em todos os instantes do indivíduo j e P_{it} é a probabilidade de acerto do item i no instante t dada as K habilidades

2.2.4 Definindo f

O que resta agora, para o fechamento do modelo, é definir f . Na verdade f é dado comparando-se várias funções através de testes, afim de obter a melhor.

O teste indicado aqui será o Akaike. Não será disposto a construção do teste, apresentado aqui de forma bem direta. O teste consiste em comparar os valores dos AIC 's, o modelo que apresentar o menor valor é o mais favorável. Define-se:

$$AIC = -2l + 2m,$$

onde l será o logaritmo natural da função de verossimilhança e m o número de parâmetros a se estimar.

De modo geral os modelos polinomiais normalmente são suficientes, contudo, em muitos casos longitudinais é previsto para f uma função sigmoide. Segue a baixo as funções sugeridas.

Tabela 2.1: Modelos Sugeridos

Polinomial $f(t|\boldsymbol{\alpha}_k) = \sum_{p=1}^P \alpha_{kp} t^{p-1}$

Gompertz $f(t|\boldsymbol{\alpha}_k) = \alpha_{k1} + \alpha_{k2} e^{-e^{-\alpha_{k3}(t-\alpha_{k4})}}$

Richards $f(t|\boldsymbol{\alpha}_k) = \alpha_{k1} + \alpha_{k2} (1 + \alpha_{k3} e^{-\alpha_{k4}(t-\alpha_{k5})})^{-\alpha_{k3}}$

Capítulo 3

Processo de Estimação

3.1 Estimação por Máxima Verossimilhança

3.1.1 Introdução

O método empregado para estimação dos parâmetros será a Estimação por Máxima Verossimilhança (EMV). Este método simples consiste em tomar a derivada da função de verossimilhança em função do parâmetro que deseja-se estimar, e igualá-la a zero afim de obter o máximo da função. Lembrando que a função de verossimilhança é a expressão que resulta na probabilidade de ocorrer a amostra, se os parâmetros forem definidos, logo, os parâmetro que maximizam a função são interpretados como os parâmetros que tornam a amostra a mais provável.

A função de verossimilhança para o parâmetro genérico w é notada como $L(w)$. Logo, faz-se:

$$\frac{\partial L(w)}{\partial w} = 0. \quad (3.1)$$

E ainda é definido como log-verossimilhança a função l dada:

$$l(w) = \ln L(w)$$

Dada uma função genérica r , note que, pela regra da cadeia:

$$\begin{aligned}\frac{\partial \ln r}{\partial w} &= \frac{\partial \ln r}{\partial r} \frac{\partial r}{\partial w} \Rightarrow \\ \frac{\partial \ln r}{\partial w} &= \frac{1}{r} \frac{\partial r}{\partial w} \Rightarrow\end{aligned}\tag{3.2}$$

$$\frac{\partial r}{\partial w} = f \frac{\partial \ln r}{\partial w}.\tag{3.3}$$

Pelo artifício apresentado na expressão 3.3 aplicado na 3.1, obtêm-se:

$$\begin{aligned}L(u) \frac{\partial l(u)}{\partial u} &= 0 \Rightarrow \\ \frac{\partial l(u)}{\partial u} &= 0.\end{aligned}\tag{3.4}$$

Logo, os parâmetros que maximizam a log-verossimilhança, também maximizam a função de verossimilhança. Por uma questão de conveniência será utilizada a log-verossimilhança. Assim, para aplicarmos o método é necessário, primeiramente, definir a tal função de log-verossimilhança.

3.1.2 Métodos da Máxima verossimilhança Marginal

Uma abordagem proposta por Bock & Lieberman (1970) separa o processo de estimação das habilidades com os demais parâmetros estabelecendo uma verossimilhança independente das habilidade: $L(\zeta, \eta)$.

Logo, é possível estimar primeiro os parâmetros dos itens e populacionais conjuntamente e depois estimar as habilidades. Este método é conhecido como Máxima Verossimilhança Marginal.

A probabilidade de um indivíduo j apresentar um determinado vetor resposta em um teste no instante t , dada a independência condicional, é:

$$P(\mathbf{U}_{jt.} | \boldsymbol{\theta}, \zeta) = \prod_{i \in I_t} P(U_{jti} | \boldsymbol{\theta}, \zeta).$$

E ainda faz-se uso da independência temporal:

$$P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) = \prod_{t=1}^T \prod_{i \in I_t} P(U_{jti} | \boldsymbol{\theta}, \zeta).\tag{3.5}$$

Pela lei da Probabilidade Total é obtido:

$$P(\mathbf{U}_{j..} | \zeta, \eta) = \int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta, \eta) g(\boldsymbol{\theta} | \eta) d\boldsymbol{\theta}$$

Note que a distribuição de $\mathbf{U}_{j..}$ não depende de $\boldsymbol{\eta}$, logo a expressão acima resumir-se-á em:

$$P(\mathbf{U}_{j..} | \zeta, \eta) = \int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \eta) d\boldsymbol{\theta}.\tag{3.6}$$

Concluindo, usa-se a independência entre as respostas de cada indivíduo, construindo assim uma função de verossimilhança marginal:

$$L(\zeta, \eta) = P(\mathbf{U} \dots | \zeta, \eta) = \prod_{j=1}^J P(\mathbf{U}_{j..} | \zeta, \eta). \quad (3.7)$$

Por fim a log-verossimilhança será:

$$l(\zeta, \eta) = \sum_{j=1}^J \ln P(\mathbf{U}_{j..} | \zeta, \eta). \quad (3.8)$$

3.2 Estimação das Habilidades

Como já mencionado, dado o método da máxima verossimilhança marginal, pode-se estimar as habilidades já dado ζ e η . Para tanto será empregado o estimador de Bayes. Pelo Teorema de Bayes escreve-se:

$$P(\boldsymbol{\theta} | \mathbf{U}_{j..}, \zeta, \eta) = \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta, \eta) g(\boldsymbol{\theta} | \zeta, \eta)}{\int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta, \eta) g(\boldsymbol{\theta} | \zeta, \eta) d\boldsymbol{\theta}}.$$

Note que a distribuição de $\mathbf{U}_{j..}$ não depende de η , assim como a distribuição de $\boldsymbol{\theta}$ não depende de ζ ; logo a expressão acima resumir-se-á em:

$$P(\boldsymbol{\theta} | \mathbf{U}_{j..}, \zeta, \eta) = \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \zeta, \eta)}{\int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \zeta, \eta) d\boldsymbol{\theta}}.$$

O estimador de Bayes é dado pela $E(\boldsymbol{\theta} | \mathbf{U}_{j..}, \zeta, \eta)$, logo conclui-se:

$$\hat{\boldsymbol{\theta}}_{.j} = \frac{\int_{\mathbb{R}^{KT}} \boldsymbol{\theta} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \zeta, \eta) d\boldsymbol{\theta}}{\int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \zeta, \eta) d\boldsymbol{\theta}}. \quad (3.9)$$

3.3 Raízes da log-verossimilhança

3.3.1 Método de Newton-Raphson

Este é um método numérico para calcular uma raiz de uma função genérica r . Baseia-se na ideia de que uma reta tangente a função em um ponto arbitrário toca o eixo das abcissas gerando uma abscissa mais próxima de uma raiz de r , com esta abscissa toma-se um novo ponto em r assim o processo é aplicado novamente, até obter-se um erro aceitável. A Figura (3.1) torna a ideia do método mais clara.

Portanto a n -ésima interação é dada:

$$x_n = x_{n-1} - \frac{r(x_{n-1})}{r'(x_{n-1})}$$

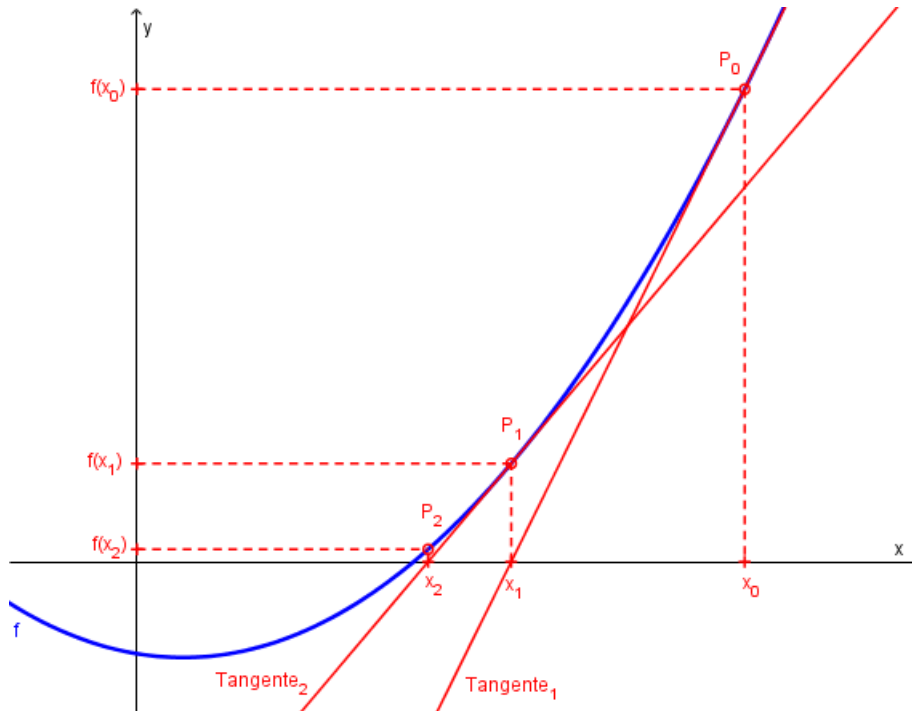


Figura 3.1: Método de Newton-Raphson

Não será preocupação do trabalho a generalização do método e a prova de sua convergência.

No caso, espera-se encontrar a raiz para a $l'(\mathbf{w})$, sendo \mathbf{w} um vetor de parâmetros genérico $\mathbf{w} = (w_1, \dots, w_N)'$, e para evitar ambiguidades o vetor de parâmetros na n -ésima iteração do método de Newton-Raphson será notada $\mathbf{w}^{(n)}$. Portanto:

$$\mathbf{w}^{(n)} = \mathbf{w}^{(n-1)} - H(\mathbf{w}^{(n-1)})^{-1}h(\mathbf{w}^{(n-1)}), \quad (3.10)$$

onde:

$$h(\mathbf{w}) = \frac{\partial l(\mathbf{w})}{\partial \mathbf{w}}$$

e

$$H(\mathbf{w}) = \frac{\partial^2 l(\mathbf{w})}{\partial \mathbf{w} \partial \mathbf{w}'},$$

portanto será notado:

$$h(w_a) = \frac{\partial l(\mathbf{w})}{\partial w_a}$$

e

$$H(w_a, w_b) = \frac{\partial^2 l(\mathbf{w})}{\partial w_a \partial w_b}.$$

Como já dito, o Capítulo 4 será responsável por desenvolver as expressões referentes a $h(\mathbf{w})$ e $H(\mathbf{w})$.

3.3.2 Método de Bock & Aitkin

Um abordagem facilitadora foi proposta por Bock & Aitkin (1981), considerando ainda o vetor de parâmetros genérico w , se w_a e w_b serem independentes for uma suposição válida, então considera-se:

$$H(w_a, w_b) = 0,$$

o que traz uma maior leveza computacional.

Com tal artifício, toma-se:

- independência entre os itens:

$$H(\zeta_a, \zeta_b) = 0, a \neq b.$$

- independência entre itens e os parâmetros populacionais:

$$H(\zeta, \eta) = 0.$$

Capítulo 4

Expressões para a estimação dos parâmetros

Este capítulo dedica-se a apresentação das expressões necessárias para a aplicação do método de Newton-Raphson. Em outras palavras, trata da exibição da primeira e segunda derivada da função de log-verossimilhança em relação aos parâmetros que deseja-se estimar, excetuando as habilidades.

4.1 Parâmetros dos itens

Nesta seção serão dadas as derivadas, primeira e segunda, da função de log-verossimilhança em relação aos parâmetros dos itens. Com essas expressões pode-se estimar os parâmetros dos itens dado os parâmetros populacionais, caso os últimos sejam desconhecidos será necessário as expressões da próxima seção para realizar a chamada estimação em bloco (Método de Bock & Aitkin).

4.1.1 1ª Derivada

Esta subseção terá como foco a primeira derivada da função de log-verossimilhança em relação aos parâmetros dos itens.

Partindo da expressão (3.8) aplica-se a derivada:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} = \sum_{j=1}^J \frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta});$$

pelo artifício apresentado na expressão (3.2) segue:

$$\frac{\partial l(\zeta, \eta)}{\partial \zeta_i} = \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..}|\zeta, \eta)} \frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\zeta, \eta). \quad (4.1)$$

O foco algébrico passa a estar apenas em $\frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\zeta, \eta)$, assim, dado a expressão (3.6), segue:

$$\begin{aligned} \frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\zeta, \eta) &= \frac{\partial}{\partial \zeta_i} \int_{\mathbb{R}^{KT}} P(\mathbf{U}_{j..}|\theta, \zeta) g(\theta|\eta) d\theta \Rightarrow \\ \frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\zeta, \eta) &= \int_{\mathbb{R}^{KT}} g(\theta|\eta) \frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\theta, \zeta) d\theta, \end{aligned}$$

portanto, o foco algébrico passa agora para $\frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\theta, \zeta)$ e com a expressão (3.3) segue:

$$\frac{\partial}{\partial \zeta_i} P(\mathbf{U}_{j..}|\theta, \zeta) = P(\mathbf{U}_{j..}|\theta, \zeta) \frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta),$$

e mais uma vez o foco é repassado, agora para $\frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta)$, com a expressão (3.5) tem-se:

$$\begin{aligned} \frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta) &= \frac{\partial}{\partial \zeta_i} \ln \prod_{t=1}^T \prod_{i^* \in I_t} P(U_{ji^*t}|\theta, \zeta) \Rightarrow \\ \frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta) &= \sum_{t=1}^T \sum_{i^* \in I_t} \frac{\partial}{\partial \zeta_i} \ln P(U_{ji^*t}|\theta, \zeta) \Rightarrow \\ \frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta) &= \sum_{t=1}^T \sum_{i^* \in I_t} \frac{1}{P(U_{ji^*t}|\theta, \zeta)} \frac{\partial}{\partial \zeta_i} P(U_{ji^*t}|\theta, \zeta). \end{aligned}$$

Note que, dado T_i o conjunto de instantes que o item i é aplicado, na somatória mais externa da expressão acima as parcelas em que $t \notin T_i$, resultam em zero devido á derivada, logo basta escrever:

$$\frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta) = \sum_{t \in T_i} \sum_{i^* \in I_t} \frac{1}{P(U_{ji^*t}|\theta, \zeta)} \frac{\partial}{\partial \zeta_i} P(U_{ji^*t}|\theta, \zeta).$$

Por outro lado, a somatória mais interna é desnecessária, pois para todo $i \neq i^*$ tem-se $\frac{\partial}{\partial \zeta_i} P(U_{ji^*t}|\theta, \zeta) = 0$, logo:

$$\frac{\partial}{\partial \zeta_i} \ln P(\mathbf{U}_{j..}|\theta, \zeta) = \sum_{t \in T_i} \frac{1}{P(U_{jit}|\theta, \zeta)} \frac{\partial}{\partial \zeta_i} P(U_{jit}|\theta, \zeta).$$

O foco, agora, recai em $\frac{\partial}{\partial \zeta_i} P(U_{jit}|\theta, \zeta)$, logo segue:

$$\begin{aligned} \frac{\partial}{\partial \zeta_i} P(U_{jit}|\theta, \zeta) &= \frac{\partial}{\partial \zeta_i} \left[P_{it}^{U_{jit}} Q_{it}^{1-U_{jit}} \right] \Rightarrow \\ \frac{\partial}{\partial \zeta_i} P(U_{jit}|\theta, \zeta) &= U_{jit} P_{it}^{U_{jit}-1} Q_{it}^{1-U_{jit}} \left(\frac{\partial P_{it}}{\partial \zeta_i} \right) - P_{it}^{U_{jit}} (1 - U_{jit}) Q_{it}^{-U_{jit}} \left(\frac{\partial P_{it}}{\partial \zeta_i} \right) \Rightarrow \\ \frac{\partial}{\partial \zeta_i} P(U_{jit}|\theta, \zeta) &= \left[U_{jit} P_{it}^{U_{jit}-1} Q_{it}^{1-U_{jit}} - P_{it}^{U_{jit}} (1 - U_{jit}) Q_{it}^{-U_{jit}} \right] \left(\frac{\partial P_{it}}{\partial \zeta_i} \right). \end{aligned}$$

Note que se $U_{jit} = 1$ então o termo entre colchetes será 1 e se $U_{jit} = 0$ este termo será -1 , logo:

$$\frac{\partial}{\partial \zeta_i} P(U_{jit} | \boldsymbol{\theta}, \boldsymbol{\zeta}) = (-1)^{U_{jit}-1} \left(\frac{\partial P_{it}}{\partial \zeta_i} \right), \quad (4.2)$$

Com os resultados acima segue:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} = \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta} | \boldsymbol{\eta}) P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \boldsymbol{\zeta}) \left\{ \sum_{t \in T_i} \frac{1}{P(U_{jit} | \boldsymbol{\theta}, \boldsymbol{\zeta})} (-1)^{U_{jit}-1} \left(\frac{\partial P_{jit}}{\partial \zeta_i} \right) \right\} d\boldsymbol{\theta}.$$

Note que:

$$\frac{1}{P(U_{jit} | \boldsymbol{\theta}, \boldsymbol{\zeta})} (-1)^{U_{jit}-1} = \frac{(-1)^{U_{jit}-1}}{P_{it}^{U_{jit}} Q_{jit}^{1-U_{jit}}} = \begin{cases} \frac{1}{P_{it}}, & \text{se } U_{jit} = 1 \\ -\frac{1}{Q_{it}}, & \text{se } U_{jit} = 0 \end{cases}$$

Assim pode-se escrever:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} = \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta} | \boldsymbol{\eta}) P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \boldsymbol{\zeta}) \left\{ \sum_{t \in T_i} \frac{U_{jit} - P_{jit}}{P_{jit} Q_{jit}} \left(\frac{\partial P_{jit}}{\partial \zeta_i} \right) \right\} d\boldsymbol{\theta}$$

Por conveniência defini-se:

$$\begin{aligned} P_{it}^* &= \left(1 + e^{-(a_{1i}\theta_{1t} + \dots + a_{Ki}\theta_{Kt} + d_i)} \right)^{-1}, \\ Q_{it}^* &= 1 - P_{it}^*, \\ W_{it} &= \frac{P_{it}^* Q_{it}^*}{P_{it} Q_{it}}, \\ g^*(\boldsymbol{\theta}) &= P(\boldsymbol{\theta} | \mathbf{U}_{j..}, \boldsymbol{\zeta}, \boldsymbol{\eta}) = \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \boldsymbol{\zeta}) g(\boldsymbol{\theta} | \boldsymbol{\eta})}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}; \end{aligned}$$

por fim, surge a expressão:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \zeta_i} = \sum_{j=1}^J \int_{\mathbb{R}^{KT}} g^*(\boldsymbol{\theta}) \left\{ \sum_{t \in T_i} (U_{jit} - P_{it}) \frac{W_{it}}{P_{it}^* Q_{it}^*} \left(\frac{\partial P_{it}}{\partial \zeta_i} \right) \right\} d\boldsymbol{\theta}$$

De agora em diante será tomado a derivada em relação a cada parâmetro específico. Primeiro será tomado um a_{ki} . Para isso será notado $M_i = -(a_{1i}\theta_{1t} + \dots + a_{Ki}\theta_{Kt} + d_i)$, assim, dado a expressão (2.3) e as convenções de notações determinadas até aqui, tem-se:

$$\begin{aligned} \frac{\partial P_{it}}{\partial a_{ki}} &= \frac{\partial}{\partial a_{ki}} \{c_i + (1 - c_i)P_{it}^*\} \Rightarrow \\ \frac{\partial P_{it}}{\partial a_{ki}} &= (1 - c_i) \frac{\partial P_{it}^*}{\partial a_{ki}} \Rightarrow \\ \frac{\partial P_{it}}{\partial a_{ki}} &= (1 - c_i) \frac{\partial(1 + e^{M_i})^{-1}}{\partial(1 + e^{M_i})} \frac{\partial(1 + e^{M_i})}{\partial e^{M_i}} \frac{\partial e^{M_i}}{\partial M_i} \frac{\partial M_i}{\partial a_{ki}} \Rightarrow \\ \frac{\partial P_{it}}{\partial a_{ki}} &= -(1 - c_i)(1 + e^{M_i})^{-2} e^{M_i} \frac{\partial M_i}{\partial a_{ki}} \Rightarrow \\ \frac{\partial P_{it}}{\partial a_{ki}} &= -(1 - c_i)(P_{it}^*)^2 e^{M_i} \frac{\partial M_i}{\partial a_{ki}}. \end{aligned}$$

Verifica-se que $P_{it}^* e^{M_i} = 1 - P_{it}^* = Q_{it}^*$, assim:

$$\begin{aligned}\frac{\partial P_{it}}{\partial a_{ki}} &= -(1 - c_i) P_{it}^* Q_{it}^* \frac{\partial M_i}{\partial a_{ki}} \Rightarrow \\ \frac{\partial P_{it}}{\partial a_{ki}} &= (1 - c_i) P_{it}^* Q_{it}^* \theta_{kt}.\end{aligned}\quad (4.3)$$

Para d_i de modo análogo obtém-se:

$$\begin{aligned}\frac{\partial P_{it}}{\partial d_i} &= -(1 - c_i) P_{it}^* Q_{it}^* \frac{\partial M_i}{\partial d_i} \Rightarrow \\ \frac{\partial P_{it}}{\partial d_i} &= (1 - c_i) P_{it}^* Q_{it}^*.\end{aligned}\quad (4.4)$$

Já para c_i :

$$\begin{aligned}\frac{\partial P_{it}}{\partial c_i} &= \frac{\partial}{\partial c_i} \{c_i + (1 - c_i) P_{it}^*\} \Rightarrow \\ \frac{\partial P_{it}}{\partial c_i} &= Q_{it}^*.\end{aligned}\quad (4.5)$$

Assim conclui-se:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial a_{ki}} = (1 - c_i) \sum_{j=1}^J \int_{\mathbb{R}^{KT}} g^*(\boldsymbol{\theta}) \left\{ \sum_{t \in T_i} (U_{jit} - P_{it}) W_{it} \theta_{kt} \right\} d\boldsymbol{\theta}, \quad (4.6)$$

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial d_i} = (1 - c_i) \sum_{j=1}^J \int_{\mathbb{R}^{KT}} g^*(\boldsymbol{\theta}) \left\{ \sum_{t \in T_i} (U_{jit} - P_{it}) W_{it} \right\} d\boldsymbol{\theta}, \quad (4.7)$$

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial c_i} = \sum_{j=1}^J \int_{\mathbb{R}^{KT}} g^*(\boldsymbol{\theta}) \left\{ \sum_{t \in T_i} (U_{jit} - P_{it}) \frac{W_{it}}{P_{it}^*} \right\} d\boldsymbol{\theta}. \quad (4.8)$$

4.1.2 2ª Derivada

Nesta subsecção será mostrada a segunda derivada da função de log-verossimilhança em relação aos parâmetros dos itens.

Para dois parâmetros de item genéricos w_1 e w_2 e pela expressão (4.1) faz-se:

$$\begin{aligned}\frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \frac{\partial}{\partial w_2} \left[\sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial}{\partial w_1} P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta}) \right] \Rightarrow \\ \frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \sum_{j=1}^J \frac{\partial}{\partial w_2} \left[\frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial}{\partial w_1} P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta}) \right] \Rightarrow \\ \frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \sum_{j=1}^J \left\{ \frac{\partial}{\partial w_2} \left[\frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \right] \frac{\partial P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_1} + \frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial^2 P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} \right\} \Rightarrow \\ \frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \sum_{j=1}^J \left\{ \frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial^2 P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} - \right. \\ &\quad \left. - \left[\frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_1} \right] \left[\frac{1}{P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})} \frac{\partial P(\mathbf{U}_{j..} | \boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2} \right] \right\}.\end{aligned}\quad (4.9)$$

A segunda parcela é inferida dos resultados da 1ª derivada. Já na primeira parcela há uma derivada de segunda ordem tomada como foco neste momento.

Por conveniência será definido:

$$V_{jit} = \frac{(-1)^{U_{jit}+1}}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})}$$

Prevendo os próximos passos será adiantado o cálculo que segue:

$$\begin{aligned} \frac{\partial V_{jit}}{\partial w_2} &= (-1)^{U_{jit}+1} \frac{\partial}{\partial w_2} \left[\frac{1}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} \right] \Rightarrow \\ \frac{\partial V_{jit}}{\partial w_2} &= (-1)^{U_{jit}+1} \left[\frac{-1}{P^2(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} \frac{\partial P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial w_2} \right] \Rightarrow \\ \frac{\partial V_{jit}}{\partial w_2} &= -\frac{(-1)^{U_{jit}+1}}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} \left[\frac{1}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} \frac{\partial P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial w_2} \right] \Rightarrow \\ \frac{\partial V_{jit}}{\partial w_2} &= -V_{jit} \left[\frac{1}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} \frac{\partial P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial w_2} \right]. \end{aligned}$$

Substituindo o resultado da expressão (4.2) e dada a definição V_{jit} conclui-se:

$$\begin{aligned} \frac{\partial V_{jit}}{\partial w_2} &= -V_{jit} \left[\frac{1}{P(U_{jit}|\boldsymbol{\theta}, \boldsymbol{\zeta})} (-1)^{U_{jit}-1} \left(\frac{\partial P_{jit}}{\partial w_2} \right) \right] \Rightarrow \\ \frac{\partial V_{jit}}{\partial w_2} &= -V_{jit}^2 \left(\frac{\partial P_{jit}}{\partial w_2} \right). \end{aligned} \quad (4.10)$$

Voltando para a derivada de segunda ordem, pelos resultados anteriores segue:

$$\begin{aligned} \frac{\partial}{\partial w_1} P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta}) &= \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta}|\boldsymbol{\eta}) P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] d\boldsymbol{\theta} \Rightarrow \\ \frac{\partial^2 P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta}|\boldsymbol{\eta}) \frac{\partial}{\partial w_2} \left\{ P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] \right\} d\boldsymbol{\theta} \Rightarrow \\ \frac{\partial^2 P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_1 \partial w_2} &= \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta}|\boldsymbol{\eta}) \left\{ \frac{\partial P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] + \right. \\ &\quad \left. + P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) \frac{\partial}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] \right\} d\boldsymbol{\theta}. \end{aligned} \quad (4.11)$$

Portanto:

$$\begin{aligned} \frac{\partial}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] &= \sum_{t \in T_i} \frac{\partial}{\partial w_2} \left[V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] \Rightarrow \\ \frac{\partial}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] &= \sum_{t \in T_i} \left[\left(\frac{\partial V_{jit}}{\partial w_2} \right) \left(\frac{\partial P_{it}}{\partial w_1} \right) + V_{jit} \left(\frac{\partial^2 P_{it}}{\partial w_1 \partial w_2} \right) \right], \end{aligned}$$

substituindo o antecipado na expressão (4.10), segue:

$$\frac{\partial}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] = \sum_{t \in T_i} \left[-V_{jit}^2 \left(\frac{\partial P_{it}}{\partial w_2} \right) \left(\frac{\partial P_{it}}{\partial w_1} \right) + V_{jit} \left(\frac{\partial^2 P_{it}}{\partial w_1 \partial w_2} \right) \right].$$

Substituindo a expressão acima em (4.11):

$$\begin{aligned} \frac{\partial^2 P(\mathbf{U}_{j..}|\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial w_2 \partial w_1} &= \int_{\mathbb{R}^{KT}} g(\boldsymbol{\theta}|\boldsymbol{\eta}) \left\{ \frac{\partial P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta})}{\partial w_2} \left[\sum_{t \in T_i} V_{jit} \left(\frac{\partial P_{it}}{\partial w_1} \right) \right] + \right. \\ &\left. + P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \boldsymbol{\zeta}) \sum_{t \in T_i} \left[-V_{jit}^2 \left(\frac{\partial P_{it}}{\partial w_2} \right) \left(\frac{\partial P_{it}}{\partial w_1} \right) + V_{jit} \left(\frac{\partial^2 P_{it}}{\partial w_2 \partial w_1} \right) \right] \right\} d\boldsymbol{\theta}. \end{aligned} \quad (4.12)$$

De modo análogo na primeira derivada, observa-se que:

$$V_{jit} = \frac{U_{jit} - P_{it}}{P_{it} Q_{it}} = (U_{jit} - P_{it}) \frac{W_{it}}{P_{it}^* Q_{it}^*} \quad (4.13)$$

Dado os resultados anteriores nesta subseção o interesse recai agora em $\left(\frac{\partial^2 P_{it}}{\partial w_2 \partial w_1} \right)$. Faz-se w um parâmetro de item genérico, antecipando álgebras futuras, note:

$$\begin{aligned} \frac{\partial P_{it}^* Q_{it}^*}{\partial w} &= \frac{\partial [P_{it}^* - (P_{it}^*)^2]}{\partial w} \Rightarrow \\ \frac{\partial P_{it}^* Q_{it}^*}{\partial w} &= \frac{\partial [P_{it}^* - (P_{it}^*)^2]}{\partial P_{it}^*} \frac{\partial P_{it}^*}{\partial w} \Rightarrow \\ \frac{\partial P_{it}^* Q_{it}^*}{\partial w} &= (1 - 2P_{it}^*) \frac{\partial P_{it}^*}{\partial w}. \end{aligned}$$

Pode-se inferir com as expressões (4.3) e (4.4) que:

$$\begin{aligned} \frac{\partial P_{it}^*}{\partial a_{ki}} &= P_{ti}^* Q_{ti}^* \theta_{kt}, \\ \frac{\partial P_{it}^*}{\partial d_i} &= P_{ti}^* Q_{ti}^*. \end{aligned}$$

Com as expressões (4.3), (4.4), (4.5) e com os resultados acima é fácil concluir:

$$\frac{\partial^2 P_{it}}{\partial a_{ki}^2} = (1 - c_i) \theta_{kt}^2 (1 - 2P_{it}^*) P_{it}^* Q_{it}^*; \quad (4.14)$$

$$\frac{\partial^2 P_{it}}{\partial a_{k_2 i} \partial a_{k_1 i}} = (1 - c_i) \theta_{k_1 t} \theta_{k_2 t} (1 - 2P_{it}^*) P_{it}^* Q_{it}^*; \quad (4.15)$$

$$\frac{\partial^2 P_{it}}{\partial d_i \partial a_{ki}} = \frac{\partial^2 P_{it}}{\partial a_{ki} \partial d_i} = (1 - c_i) \theta_{kt} (1 - 2P_{it}^*) P_{it}^* Q_{it}^*; \quad (4.16)$$

$$\frac{\partial P_{it}}{\partial c_i \partial a_{ki}} = \frac{\partial P_{it}}{\partial a_{ki} \partial c_i} = -P_{it}^* Q_{it}^* \theta_{kt}; \quad (4.17)$$

$$\frac{\partial^2 P_{it}}{\partial d_i^2} = (1 - c_i) (1 - 2P_{it}^*) P_{it}^* Q_{it}^*; \quad (4.18)$$

$$\frac{\partial^2 P_{it}}{\partial c_i \partial d_i} = -(1 - 2P_{it}^*) P_{it}^* Q_{it}^*; \quad (4.19)$$

$$\frac{\partial^2 P_{it}}{\partial c_i^2} = 0; \quad (4.20)$$

$$\frac{\partial^2 P_{it}}{\partial d_i \partial c_i} = -P_{it}^* Q_{it}^*. \quad (4.21)$$

4.2 Parâmetros Populacionais: Sem curvas Crescimento

Esta seção tem como objetivo mostrar a primeira e segunda derivadas da função de log-verossimilhança em relação aos parâmetros populacionais, desconsiderando um modelo de curva de crescimento. Caso os parâmetros dos itens sejam dados, as expressões aqui são suficientes para estimação dos Parâmetros populacionais, caso contrário, como já dito, é necessário as expressões dos parâmetros dos itens também.

4.2.1 1ª Derivada

Nesta subseção o foco é mostrar a primeira derivada da função de log-verossimilhança em relação aos parâmetros populacionais, desconsiderando um modelo de curva de crescimento.

Pela expressão (3.8) tem-se:

$$\begin{aligned}\frac{\partial l(\zeta, \eta)}{\partial \eta} &= \frac{\partial}{\partial \eta} \left[\sum_{j=1}^J P(\mathbf{U}_{j..} | \zeta, \eta) \right] \Rightarrow \\ \frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \frac{\partial \ln P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta},\end{aligned}$$

dado o artifício da expressão (3.2) segue:

$$\frac{\partial l(\zeta, \eta)}{\partial \eta} = \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \zeta, \eta)} \frac{\partial P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta} \quad (4.22)$$

e substituindo a expressão (3.5) na expressão acima tem-se:

$$\begin{aligned}\frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \zeta, \eta)} \frac{\partial}{\partial \eta} \left[\int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \eta) d\boldsymbol{\theta} \right] \Rightarrow \\ \frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..} | \zeta, \eta)} \int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) \frac{\partial g(\boldsymbol{\theta} | \eta)}{\partial \eta} d\boldsymbol{\theta} \Rightarrow \\ \frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \int_{\mathbb{R}^{TK}} \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta)}{P(\mathbf{U}_{j..} | \zeta, \eta)} \frac{\partial g(\boldsymbol{\theta} | \eta)}{\partial \eta} d\boldsymbol{\theta} \Rightarrow \\ \frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \int_{\mathbb{R}^{TK}} \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \eta)}{P(\mathbf{U}_{j..} | \zeta, \eta)} \frac{\partial \ln g(\boldsymbol{\theta} | \eta)}{\partial \eta} d\boldsymbol{\theta},\end{aligned}$$

por conveniência de notação será tomado $g_j^*(\boldsymbol{\theta}) = \frac{P(\mathbf{U}_{j..} | \boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta} | \eta)}{P(\mathbf{U}_{j..} | \zeta, \eta)}$, concluindo:

$$\frac{\partial l(\zeta, \eta)}{\partial \eta} = \sum_{j=1}^J \int_{\mathbb{R}^{TK}} g_j^*(\boldsymbol{\theta}) \frac{\partial \ln g(\boldsymbol{\theta} | \eta)}{\partial \eta} d\boldsymbol{\theta}. \quad (4.23)$$

A distribuição convencionada para este trabalho, discutida no capítulo 2, é a normal multivaria com parâmetros $(\boldsymbol{\mu}, \boldsymbol{\Sigma})$. Logo, tomando $V = \boldsymbol{\Sigma}^{-1}$:

$$\begin{aligned}
g(\boldsymbol{\theta}|\boldsymbol{\eta}) &= \frac{1}{(2\pi)^{TK/2}|\boldsymbol{\Sigma}|^{1/2}} e^{-(\boldsymbol{\theta}-\boldsymbol{\mu})'V(\boldsymbol{\theta}-\boldsymbol{\mu})/2} \Rightarrow \\
\ln g(\boldsymbol{\theta}|\boldsymbol{\eta}) &= -\frac{TK}{2} \ln(2\pi) - \frac{1}{2} \ln|\boldsymbol{\Sigma}| - (\boldsymbol{\theta}-\boldsymbol{\mu})'V(\boldsymbol{\theta}-\boldsymbol{\mu})/2 \Rightarrow \\
\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} &= -\frac{1}{2} \left\{ \frac{\partial \ln|\boldsymbol{\Sigma}|}{\partial \boldsymbol{\eta}} + \frac{\partial(\boldsymbol{\theta}-\boldsymbol{\mu})'V(\boldsymbol{\theta}-\boldsymbol{\mu})}{\partial \boldsymbol{\eta}} \right\}. \tag{4.24}
\end{aligned}$$

Especificando a derivada para $\boldsymbol{\mu}$ segue:

$$\begin{aligned}
\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= -\frac{1}{2} \left\{ \frac{\partial(\boldsymbol{\theta}-\boldsymbol{\mu})'V(\boldsymbol{\theta}-\boldsymbol{\mu})}{\partial \boldsymbol{\mu}} \right\} \Rightarrow \\
\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V(\boldsymbol{\theta}-\boldsymbol{\mu}). \tag{4.25}
\end{aligned}$$

Para prosseguir nos cálculos é conveniente definir:

$$\begin{aligned}
\tilde{\boldsymbol{\mu}}_j &= \int_{\mathbb{R}^{TK}} \boldsymbol{\theta} g_j^*(\boldsymbol{\theta}) d\boldsymbol{\theta}, \\
\bar{\boldsymbol{\mu}} &= \frac{1}{J} \sum_{j=1}^J \tilde{\boldsymbol{\mu}}_j.
\end{aligned}$$

Substituindo a expressão (4.25) em (4.23) obtém-se:

$$\begin{aligned}
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= \sum_{j=1}^J \int_{\mathbb{R}^{TK}} g_j^*(\boldsymbol{\theta}) \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\mu}} d\boldsymbol{\theta} \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= \sum_{j=1}^J \int_{\mathbb{R}^{TK}} g_j^*(\boldsymbol{\theta}) V(\boldsymbol{\theta}-\boldsymbol{\mu}) d\boldsymbol{\theta} \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V \sum_{j=1}^J \int_{\mathbb{R}^{TK}} [\boldsymbol{\theta} g_j^*(\boldsymbol{\theta}) - \boldsymbol{\mu} g_j^*(\boldsymbol{\theta})] d\boldsymbol{\theta} \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V \sum_{j=1}^J \left[\int_{\mathbb{R}^{TK}} \boldsymbol{\theta} g_j^*(\boldsymbol{\theta}) d\boldsymbol{\theta} - \boldsymbol{\mu} \int_{\mathbb{R}^{TK}} g_j^*(\boldsymbol{\theta}) d\boldsymbol{\theta} \right] \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V \sum_{j=1}^J [\tilde{\boldsymbol{\mu}}_j - \boldsymbol{\mu}] \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V \left[\sum_{j=1}^J \tilde{\boldsymbol{\mu}}_j - \boldsymbol{\mu} \sum_{j=1}^J 1 \right] \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V \left[\sum_{j=1}^J \tilde{\boldsymbol{\mu}}_j - J\boldsymbol{\mu} \right] \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V [J\bar{\boldsymbol{\mu}} - J\boldsymbol{\mu}] \Rightarrow \\
\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\mu}} &= V J(\bar{\boldsymbol{\mu}} - \boldsymbol{\mu}). \tag{4.26}
\end{aligned}$$

Agora nos resta especificar a derivada para os parâmetros referentes a estrutura de covariância: ρ_1 e ρ_2 . Para tanto, é necessário, primeiramente, expor algumas propriedades do produto de Kronecker aplicadas as matrizes de covariâncias já mencionadas:

- $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$
- $|A \otimes B| = |A|^T |B|^K$, lembrando que T é o tamanho da matriz B enquanto K é o tamanho da matriz A .

Tomando a expressão (4.24) e especificando a derivada para um ρ_i , segue:

$$\begin{aligned} \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_i} &= -\frac{1}{2} \left\{ \frac{\partial \ln |\boldsymbol{\Sigma}|}{\partial \rho_i} + \frac{\partial(\boldsymbol{\theta} - \boldsymbol{\mu})' V(\boldsymbol{\theta} - \boldsymbol{\mu})}{\partial \rho_i} \right\} \Rightarrow \\ \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_i} &= -\frac{1}{2} \left\{ \frac{\partial \ln |A \otimes B|}{\partial \rho_i} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \frac{\partial V}{\partial \rho_i} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\} \Rightarrow \\ \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_i} &= -\frac{1}{2} \left\{ \frac{\partial \ln |A|^T |B|^K}{\partial \rho_i} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \frac{\partial V}{\partial \rho_i} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\} \Rightarrow \\ \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_i} &= -\frac{1}{2} \left\{ T \frac{\partial \ln |A|}{\partial \rho_i} + K \frac{\partial \ln |B|}{\partial \rho_i} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \frac{\partial(A^{-1} \otimes B^{-1})}{\partial \rho_i} (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}. \end{aligned}$$

Especificando a derivada para ρ_1 , partindo da expressão acima:

$$\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_1} = -\frac{1}{2} \left\{ T \frac{\partial \ln |A|}{\partial \rho_1} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \left[\frac{\partial A^{-1}}{\partial \rho_1} \otimes B^{-1} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}. \quad (4.27)$$

Agora especificando a derivada para ρ_2 , partindo da mesma expressão:

$$\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_2} = -\frac{1}{2} \left\{ K \frac{\partial \ln |B|}{\partial \rho_2} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \left[A^{-1} \otimes \frac{\partial B^{-1}}{\partial \rho_2} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}. \quad (4.28)$$

Com as expressões (4.23), (4.27) e (4.28) conclui-se a primeira derivada em relação aos parâmetros de correlação.

4.2.2 2ª Derivada

Nesta subsecção o foco é mostrar a segunda derivada da função de log-verossimilhança em relação aos parâmetros populacionais, desconsiderando algum modelo de curva de crescimento.

Partindo da expressão (4.22) segue:

$$\begin{aligned}
\frac{\partial l(\zeta, \eta)}{\partial \eta} &= \sum_{j=1}^J \frac{1}{P(\mathbf{U}_{j..}|\zeta, \eta)} \frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta} \Rightarrow \\
\frac{\partial^2 l(\zeta, \eta)}{\partial \eta \partial \eta'} &= \sum_{j=1}^J \frac{\partial}{\partial \eta'} \left[\frac{1}{P(\mathbf{U}_{j..}|\zeta, \eta)} \frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta} \right] \Rightarrow \\
\frac{\partial^2 l(\zeta, \eta)}{\partial \eta \partial \eta'} &= \sum_{j=1}^J \left[\frac{\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'}}{P(\mathbf{U}_{j..}|\zeta, \eta)} - \frac{\frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta'}}{P^2(\mathbf{U}_{j..}|\zeta, \eta)} \frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta} \right] \Rightarrow \\
\frac{\partial^2 l(\zeta, \eta)}{\partial \eta \partial \eta'} &= \sum_{j=1}^J \left[\frac{\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'}}{P(\mathbf{U}_{j..}|\zeta, \eta)} - \left(\frac{\frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta}}{P(\mathbf{U}_{j..}|\zeta, \eta)} \right) \left(\frac{\frac{\partial P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta}}{P(\mathbf{U}_{j..}|\zeta, \eta)} \right)' \right] \quad (4.29)
\end{aligned}$$

Portanto, para finalizar só resta calcular $\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'}$, assim dado a expressão (3.6) segue:

$$\begin{aligned}
\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'} &= \frac{\partial}{\partial \eta'} \left\{ \frac{\partial}{\partial \eta} \left[\int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta}|\eta) d\boldsymbol{\theta} \right] \right\} \Rightarrow \\
\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'} &= \frac{\partial}{\partial \eta'} \left\{ \int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta}|\eta) \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta} d\boldsymbol{\theta} \right\} \Rightarrow \\
\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'} &= \int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \zeta) \left\{ g(\boldsymbol{\theta}|\eta) \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta'} \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta} + g(\boldsymbol{\theta}|\eta) \frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta \partial \eta'} \right\} d\boldsymbol{\theta} \Rightarrow \\
\frac{\partial^2 P(\mathbf{U}_{j..}|\zeta, \eta)}{\partial \eta \partial \eta'} &= \int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..}|\boldsymbol{\theta}, \zeta) g(\boldsymbol{\theta}|\eta) \left\{ \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta'} \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta} + \frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta \partial \eta'} \right\} d\boldsymbol{\theta} \quad (4.30)
\end{aligned}$$

Com os resultados anteriores, o problema resume-se em encontrar $\frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \eta \partial \eta'}$. Especificando para parâmetro $\boldsymbol{\mu}$ e com o resultado da expressão (4.26) segue:

$$\frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \boldsymbol{\mu} \partial \boldsymbol{\mu}'} = -V. \quad (4.31)$$

Ainda com a expressão (4.26) e com as expressões (4.27) e (4.28) é imediato o seguinte resultado:

$$\frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \boldsymbol{\mu} \partial \rho_i} = \frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \rho_i \partial \boldsymbol{\mu}} = \frac{\partial V}{\partial \rho_i} (\boldsymbol{\theta} - \boldsymbol{\mu}). \quad (4.32)$$

concluindo:

$$\frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \boldsymbol{\mu} \partial \rho_1} = \left[\frac{\partial A^{-1}}{\partial \rho_1} \otimes B^{-1} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}), \quad (4.33)$$

$$\frac{\partial \ln g(\boldsymbol{\theta}|\eta)}{\partial \boldsymbol{\mu} \partial \rho_2} = \left[A^{-1} \otimes \frac{\partial B^{-1}}{\partial \rho_2} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}). \quad (4.34)$$

Já apenas com as expressões (4.27) e (4.28) é fácil concluir:

$$\frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \rho_1 \partial \rho_2} = \frac{\partial^2 \ln g(\boldsymbol{\theta}|\eta)}{\partial \rho_2 \partial \rho_1} = -\frac{1}{2} \left\{ (\boldsymbol{\theta} - \boldsymbol{\mu})' \left[\frac{\partial A^{-1}}{\partial \rho_1} \otimes \frac{\partial B^{-1}}{\partial \rho_2} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}. \quad (4.35)$$

e ainda:

$$\frac{\partial^2 \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_1^2} = -\frac{1}{2} \left\{ T \frac{\partial^2 \ln|A|}{\partial \rho_1^2} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \left[\frac{\partial^2 A^{-1}}{\partial \rho_1^2} \otimes B^{-1} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}, \quad (4.36)$$

$$\frac{\partial^2 \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \rho_2^2} = -\frac{1}{2} \left\{ K \frac{\partial^2 \ln|B|}{\partial \rho_2^2} + (\boldsymbol{\theta} - \boldsymbol{\mu})' \left[A^{-1} \otimes \frac{\partial^2 B^{-1}}{\partial \rho_2^2} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}) \right\}. \quad (4.37)$$

Assim torna-se finalizada a segunda derivada com seu resultado nas expressões de (4.30), (4.27), (4.28), (4.32), (4.33), (4.34), (4.35), (4.36) e (4.37).

4.3 Parâmetros Populacionais: Com curvas Crescimento

O objetivo, agora, consiste em deduzir as expressões para a estimação do vetor de parâmetros $\boldsymbol{\alpha}$, assim como as novas expressões para a estimação dos parâmetros ρ_1 e ρ_2 , pois estes são algebricamente dependentes do primeiro. Ou seja, tomando $\boldsymbol{\eta} = (\boldsymbol{\alpha}', \rho_1, \rho_2)'$, quer-se $\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}}$ e $\frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta} \partial \boldsymbol{\eta}'}$.

4.3.1 1ª Derivada

Nesta subseção o foco é mostrar a primeira derivada da função de log-verosimilhança em relação aos parâmetros populacionais. Lembrando que para os parâmetros ρ_1 e ρ_2 ainda valem as mesmas expressões calculadas na seção anterior.

Pode-se reutilizar as demonstrações no caso “sem curva de crescimento” até a expressão (4.23). Assim o ponto de partida é:

$$\frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \boldsymbol{\eta}} = \sum_{j=1}^J \int_{\mathbb{R}^{TK}} g_j^*(\boldsymbol{\theta}) \frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}} d\boldsymbol{\theta},$$

portanto o foco recai para $\frac{\partial \ln g(\boldsymbol{\theta}|\boldsymbol{\eta})}{\partial \boldsymbol{\eta}}$.

Para prosseguir é conveniente definir:

$$\boldsymbol{\mu}_{pk}^{(1)} = \frac{\partial \boldsymbol{\mu}}{\partial \alpha_{pk}}.$$

Partindo da expressão (4.26) e dada a regra da cadeia conclui-se:

$$\begin{aligned} \frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \alpha_{pk}} &= \left(\frac{\partial \boldsymbol{\mu}}{\partial \alpha_{pk}} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}). \Rightarrow \\ \frac{\partial l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \alpha_{pk}} &= \left(\boldsymbol{\mu}_{pk}^{(1)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}). \end{aligned} \quad (4.38)$$

4.3.2 2ª Derivada

Nesta subseção o foco é mostrar a segunda derivada da função de log-verosimilhança em relação aos parâmetros populacionais.

Da mesma forma que na primeira, pode-se reutilizar as demonstrações anteriores até a expressão (4.29):

$$\frac{\partial^2 l(\zeta, \eta)}{\partial \eta \partial \eta'} = \sum_{j=1}^J \left[\frac{\frac{\partial^2 P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta \partial \eta'}}{P(\mathbf{U}_{j..} | \zeta, \eta)} - \left(\frac{\frac{\partial P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta}}{P(\mathbf{U}_{j..} | \zeta, \eta)} \right) \left(\frac{\frac{\partial P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta}}{P(\mathbf{U}_{j..} | \zeta, \eta)} \right)' \right],$$

assim o foco recai sobre $\frac{\partial^2 P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta \partial \eta'}$ e segundo expressão (4.30):

$$\frac{\partial^2 P(\mathbf{U}_{j..} | \zeta, \eta)}{\partial \eta \partial \eta'} = \int_{\mathbb{R}^{TK}} P(\mathbf{U}_{j..} | \theta, \zeta) g(\theta | \eta) \left\{ \frac{\partial \ln g(\theta | \eta)}{\partial \eta'} \frac{\partial \ln g(\theta | \eta)}{\partial \eta} + \frac{\partial^2 \ln g(\theta | \eta)}{\partial \eta \partial \eta'} \right\} d\theta,$$

e assim, finalmente o problema resume-se em resolver $\frac{\partial^2 \ln g(\theta | \eta)}{\partial \eta \partial \eta'}$. Para prosseguir é conveniente definir:

$$\boldsymbol{\mu}_{p_1 k_1, p_2 k_2}^{(2)} = \frac{\partial^2 \boldsymbol{\mu}}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}}.$$

Assim, segue a partir da expressão (4.38):

$$\begin{aligned} \frac{\partial l(\zeta, \eta)}{\partial \alpha_{p_1 k_1}} &= \left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}) \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}} &= \frac{\partial}{\partial \alpha_{p_2 k_2}} \left[\left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}) \right] \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}} &= \frac{\partial}{\partial \alpha_{p_2 k_2}} \left[\left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V\boldsymbol{\theta} - \left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V\boldsymbol{\mu} \right] \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}} &= \frac{\partial}{\partial \alpha_{p_2 k_2}} \left[\left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V\boldsymbol{\theta} - \left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V\boldsymbol{\mu} \right] \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}} &= \left(\boldsymbol{\mu}_{p_1 k_1, p_2 k_2}^{(2)} \right)' V\boldsymbol{\theta} - \left(\boldsymbol{\mu}_{p_1 k_1, p_2 k_2}^{(2)} \right)' V\boldsymbol{\mu} - \left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V \left(\boldsymbol{\mu}_{p_2 k_2}^{(1)} \right) \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{p_1 k_1} \partial \alpha_{p_2 k_2}} &= \left(\boldsymbol{\mu}_{p_1 k_1, p_2 k_2}^{(2)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}) - \left(\boldsymbol{\mu}_{p_1 k_1}^{(1)} \right)' V \left(\boldsymbol{\mu}_{p_2 k_2}^{(1)} \right) \end{aligned} \quad (4.39)$$

É conveniente notar que se $k_1 \neq k_2$ então a expressão acima é nula.

Ainda partindo da expressão (4.38) segue:

$$\begin{aligned} \frac{\partial l(\zeta, \eta)}{\partial \alpha_{pk}} &= \left(\boldsymbol{\mu}_{pk}^{(1)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}) \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{pk} \partial \rho_i} &= \frac{\partial}{\partial \rho_i} \left[\left(\boldsymbol{\mu}_{pk}^{(1)} \right)' V(\boldsymbol{\theta} - \boldsymbol{\mu}) \right] \Rightarrow \\ \frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{pk} \partial \rho_i} &= \left(\boldsymbol{\mu}_{pk}^{(1)} \right)' \frac{\partial V}{\partial \rho_i} (\boldsymbol{\theta} - \boldsymbol{\mu}), \end{aligned}$$

logo, conclui-se:

$$\frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{pk} \partial \rho_1} = \left(\boldsymbol{\mu}_{pk}^{(1)} \right)' \left[\frac{\partial A^{1-}}{\partial \rho_1} \otimes B^{-1} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}), \quad (4.40)$$

$$\frac{\partial^2 l(\zeta, \eta)}{\partial \alpha_{pk} \partial \rho_2} = \left(\boldsymbol{\mu}_{pk}^{(1)} \right)' \left[A^{1-} \otimes \frac{\partial B^{-1}}{\partial \rho_2} \right] (\boldsymbol{\theta} - \boldsymbol{\mu}). \quad (4.41)$$

E para finalizar, com a expressão 4.32 é fácil concluir:

$$\frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \alpha_{pk} \partial \rho_i} = \frac{\partial^2 l(\boldsymbol{\zeta}, \boldsymbol{\eta})}{\partial \rho_i \partial \alpha_{pk}} \quad (4.42)$$

Concluindo, os resultados desta subseção estão nas expressões (4.29), (4.30), (4.40), (4.40), (4.41) e (4.42).

Capítulo 5

Considerações Finais

Dado o modelo multidimensional com análise longitudinal com curva de crescimento e as expressões para a aplicação do método de Newton-Raphson como resultados desta monografia, o esperado é a continuação do trabalho através da implementação computacional e por fim a análise do modelo dentro de aplicações reais. Outra perspectiva é a expansão do modelo, principalmente no que diz respeito a utilização de outras matrizes de covariância, assim como a análise da eficiência da matriz apresentada, dentro das aplicações.

Também pode-se imaginar, além do método de Newton-Raphson para a estimação dos parâmetros, outros métodos como "Scoring" de Fisher e Algoritmo EM. Note que mesmo utilizando outros métodos muitas vezes as expressões apresentadas aqui se fazem necessárias.

Outro aspecto da implementação é resolução das integrais, uma técnica numérica simples é a lei do trapézio, porém deve-se pensar também na velocidade computacional, por tal motivo a investigação de outros métodos torna-se necessária como o Método de Monte Carlo e suas variantes. Perceba que é possível, sempre que conveniente pra o implementador, utilizar métodos numéricos para a resolução também de derivadas.

Referências Bibliográficas

- [1] Akaike , H. A new look at the statistical model identification. IEEE Transactions on Automatic Control., Boston, v.19, n.6, p.716-723, Dec. 1974.
- [2] Andrade, D.F., Tavares, H.R. e Valle, R.C. (2000). Teoria da Resposta ao Item: Conceitos e Aplicações. Associação Brasileira de Estatística: São Paulo.
- [3] Bock, R. D.; Gibbons, R.; Muraki, E. Full-information item factor analysis. Psychological Measurement, v. 12, n. 3, p. 261 – 281, 1988.
- [4] Birnbaum, A. (1968). Some Latent Trait Models and Their Use in Infering an Examinee's Ability. In F.M. Lord & M.R. Novick. Statistical Theories of Mental Test Scores. Reading, MA:Addison-Wesley.
- [5] Fragoso, T. M. (2010). Modelos Multidimensionais da Teoria de Resposta ao Item. Tiago de Miranda Fragoso. São Carlos, SP: USP, ICMC, 20110.
- [6] Lord, F. M. (1952). A theory of test scores (No. 7). Psychometric Monograph.
- [7] Rasch, G. (1960). Probabilistic Models for Some Intelligence and Attainment Tests. Copenhagen : Danish Institute for Educational Research.
- [8] Reckase, M. D. Multidimensional Item Response Theory. USA: Springer, 2009.
- [9] Tavares, H.R. Teoria da Resposta ao Item para Dados Longitudinais. 2001. 118 f. Tese (Doutorado) - USP, São Paulo, 2001.

