Zusatzmaterial zu Kap. 19: Rechenbeispiele zur Parameterschätzung und Messgenauigkeit in der Item-Response-Theorie (IRT) (Norman Rose)

Beispiel 1: Eine Forscherin/ein Forscher sind an der Verteilung eines Merkmals interessiert, dass sie anhand eines IRT basierten Testverfahrens erhoben haben. Dabei ist vor allem die Varianz des Merkmals von Interesse. Die marginale Reliabilität sei .70. Die Streuung der EAP-Schätzer sei $s =$ 1.20. Welche Werte für die Varianzen erwarten Sie für die übrigen Personenparameterschätzer (ML-, gewichteter ML-, MAP- und PV-Schätzer), und wie sind die Unterschiede zu erklären?

Aus Gleichung 19.111 folgt, dass die Varianz der latenten Variable θ geschätzt kann, indem man die Varianz $s^2(EAP) = 1.20$ durch die marginale Reliabilität $Rel = .70$ teilt. Dies ergibt einen Wert von $Var(\theta) = 1.71$. Die wahre Varianz der latenten Variablen ist somit größer als die Varianz der EAP-Schätzer. Dies ist auf den Shrinkage-Effekt zurückzuführen, von dem auch die MAP-Schätzer betroffen sind. Es ist zu erwarten, dass die Varianz der MAP-Schätzer ebenfalls bei rund 1.20 liegt. Sind viele individuelle A-posteriori-Verteilung extrem schief verteilt, so weisen EAP-Schätzer mitunter eine etwas größere Varianz als MAP Schätzer auf. Die Varianz des ML-Schätzer ist additiv aus der Varianz der latenten Variable und Varianz des Schätzfehlers zusammengesetzt. Daher ist die Varianz der ML-Schätzer nicht nur größer als die der EAP und MAP Schätzer, sondern auch größer als die Varianz der latenten Variablen θ. Im vorliegenden Beispiel lässt sich der Wert auch genauer abschätzen, indem man den oben geschätzten Wert $\mathrm{Var}(\theta) = 1.71$ durch die marginale Reliabilität von 0.70 teilt. Dadurch ergibt sich ein geschätzter Werte der Varianz der ML-Schätzer von 2.45. Der gewichtete ML-Schätzer ist dem ML-Schätzer sehr ähnlich, hat aber eine etwas kleinere Varianz als der ML-Schätzer und dennoch eine größere Varianz als die latenten Variablen. Die Varianz der PVs ist dagegen eine unverfälschter Schätzer der Varianz der latenten Variablen, sodass $Var(\theta) \approx Var(\text{PVs})$. Würde die Zahl der genierten PVs gegen unendlich gehen, so konvergiert die Verteilung der PVs gegen die wahre Verteilung von θ, sodass gilt $Var(\theta) = Var(\text{PVs})$.

Beispiel 2: Die vorgestellten Schätzverfahren sind letztlich Anwendungen von Mathematik, die zu einem guten Teil bereits in der Sekundarstufe 2 behandelt werden. So können die im Text gegebenen ersten und zweiten partiellen Ableitungen der logarithmierten Likelihood oder der A-priori-Verteilungen nach den Parametern selbst bestimmt werden. Dabei müssen die entsprechenden Rechenregeln für die Differentialrechnung angewendet werden. Bilden Sie bitte die folgenden Ableitungen:

a. Erste partielle Ableitung der logarithmierten Antwortmusterwahrscheinlichkeit (s. Gleichung 19.81) nach dem Personenparameter θ_v.
b. Erste partielle Ableitung des natürlichen Logarithmus der Normalverteilung der latenten Variablen $\theta \sim N(\mu_\theta, \sigma_\theta)$, nach der Personenvariablen θ.

2a. Die Ableitung der log. Antwortmusterwahrscheinlichkeit bezüglich der latenten Variablen wird hier ausgehend von Gleichung 19.15 gezeigt. Die Ableitung lässt sich wie folgt schreiben:

$$
\begin{aligned}
\frac{d}{d\theta_v} l(\theta_v) &= \sum_{i=1}^{k} \ln\left[P(Y_{vi}=1 \mid \theta_v; \boldsymbol{\iota}_i)^{y_{vi}} \, P(Y_{vi}=0 \mid \theta_v; \boldsymbol{\iota}_i)^{(1-y_{vi})} \right] \\
&= \sum_{i=1}^{k} \left[y_{vi} \ln P(Y_{vi}=1 \mid \theta_v; \boldsymbol{\iota}_i) + (1-y_{vi}) \ln P(Y_{vi}=0 \mid \theta_v; \boldsymbol{\iota}_i) \right] \\
&= \sum_{i=1}^{k} \left[y_{vi} \frac{d}{d\theta_v} \ln P(Y_{vi}=1 \mid \theta_v; \boldsymbol{\iota}_i) + (1-y_{vi}) \frac{d}{d\theta_v} \ln P(Y_{vi}=0 \mid \theta_v; \boldsymbol{\iota}_i) \right]
\end{aligned}
$$

In Zeile 3 der Gleichung wird von der Summenregel Gebrauch gemacht, der zufolge die Ableitung einer Summe gleich der Summe der Ableitungen der einzelnen Summanden ist. Die Terme y_{vi} und $(1 - y_{vi})$ sind Konstanten, die bei der partiellen Ableitung vor die Ableitung gezogen werden können.

Nun müssen zwei Ableitungen gebildet werden: erstens die erste partielle Ableitung des Logarithmus für die Lösungswahrscheinlichkeit bei gegebener latenter Variablen und den Itemparametern:

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=1|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{d}{d\theta_v}\ln\left(\frac{1}{1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}\right)$$

und zweitens die Gegenwahrscheinlichkeit:

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=0|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{d}{d\theta_v}\ln\left(\frac{1}{1+\exp\left[\alpha_i\left(\theta_v-\beta_i\right)\right]}\right)$$

Da für die erste Ableitung des Logarithmus ln[$f(X)$] einer Funktion $f(X)$ allgemein gilt:

$$\frac{d}{d\theta_v}\ln\left[f\left(X\right)\right]=\frac{f'\left(X\right)}{f\left(X\right)},$$

folgt, dass

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=1|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{1}{P\left(Y_{vi}=1|\,\theta_v;\boldsymbol{\iota}_i\right)}\frac{d}{d\theta_v}\frac{1}{1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}$$

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=0|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{1}{P\left(Y_{vi}=0|\,\theta_v;\boldsymbol{\iota}_i\right)}\frac{d}{d\theta_v}\frac{1}{1+\exp\left[\alpha_i\left(\theta_v-\beta_i\right)\right]}$$

Nun werden auf der rechten Seite der beiden Gleichungen die Ableitungen der Modellgleichung für die bedingten Wahrscheinlichkeiten $P(Y_{vi}=1\mid\theta_v;\ \iota_i)$ und $P(Y_{vi}=0\mid\theta_v;\ \iota_i)$ benötigt. Da es sich bei den Modellgleichungen um verkette Funktionen der Personenvariable handelt, müssen die Ableitungen unter Verwendung der Kettenregel erfolgen. Diese besagt allgemein, dass die Ableitung einer verketten Funktion gleich dem Produkt der Ableitung der „äußeren“ Funktion und der Ableitung der „inneren“ Funktion ist. Im vorliegenden Fall ist die innere Funktion $k(\theta_v) = 1 + \exp[-\alpha_i(\theta_v - \beta_i)]$ und die äußere Funktion $g(\theta_v) = k(\theta_v)^{-1}$, sodass:

$$\frac{d}{d\theta_v}P\left(Y_{vi}=1|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{d}{dk\left(\theta_v\right)}g\left[k\left(\theta_v\right)\right]\frac{d}{d\theta_v}k\left(\theta_v\right)$$

Die erste partielle Ableitung der inneren Funktion ist

$$\begin{aligned}\frac{d}{d\theta_v}k\left(\theta_v\right)&=\frac{d}{d\theta_v}\left(1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]\right)\\&=0+\frac{d}{d\theta_v}\exp\left[-\alpha_i\theta_v+\alpha_i\beta_i\right]\\&=-\alpha_i\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]\end{aligned}$$

Dabei werden in Zeile 2 die Summenregel und in Zeile 3 die Rechenregel für die Ableitung von exponenzierten Funktionen exp[$f(X)$] verwendet, die in allgemeiner Form lautet:

$$\frac{d}{dX}\exp\left[f\left(X\right)\right]=f'\left(X\right)\exp\left[f\left(X\right)\right]$$

Die erste partielle Ableitung der äußeren Funktion ist:

$$\begin{aligned}\frac{d}{dk\left(\theta_v\right)}g\left[k\left(\theta_v\right)\right]&=\frac{d}{d\theta_v}k\left(\theta_v\right)^{-1}\\&=-k\left(\theta_v\right)^{-2}\\&=-\frac{1}{\left(1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]\right)^2}\end{aligned}$$

Das Produkt der Ableitungen von innerer und äußerer Funktion ist schließlich:

$$\begin{aligned}\frac{d}{d\theta_v}P\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)&=\frac{\alpha_i\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}{\left(1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]\right)^2}\\&=\alpha_i\frac{1}{1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}\frac{\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}{1+\exp\left[-\alpha_i\left(\theta_v-\beta_i\right)\right]}\\&=\alpha_iP\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)\end{aligned}$$

Unter analogem Vorgehen ergibt sich die erste Ableitung für modellimplizierte bedingte Wahrscheinlichkeit das Item i nicht zu lösen. Der resultierende Term unterscheidet sich lediglich im Vorzeichen von der vorangegangenen Gleichung:

$$\frac{d}{d\theta_v}P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)=-\alpha_iP\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)$$

Die ersten Ableitungen der Modellgleichungen für Lösen und Nichtlösen des Items i können nun in die entsprechenden Gleichungen der ersten Ableitungen der logarithmierten Modellgleichungen für Lösen und Nichtlösen des Items i eingesetzt werden (s.o.).

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{\alpha_iP\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)}{P\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)}$$

$$\frac{d}{d\theta_v}\ln P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)=\frac{-\alpha_iP\left(Y_{vi}=1\,|\,\theta_v;\boldsymbol{\iota}_i\right)P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)}{P\left(Y_{vi}=0\,|\,\theta_v;\boldsymbol{\iota}_i\right)}$$

Setzt man diese beiden Terme in die Ausgangsgleichung der Log-Likelihood-Funktion für ein Antwortmuster ein, ergibt sich durch umstellen und vereinfachen:

$$\begin{aligned}\frac{dl(\theta_v)}{d\theta_v} &= \sum_{i=1}^{k}\left[y_{vi}\frac{d}{d\theta_v}\ln P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)+(1-y_{vi})\frac{d}{d\theta_v}\ln P(Y_{vi}=0|\theta_v;\boldsymbol{\iota}_i)\right]\\
&= \sum_{i=1}^{k}\left[y_{vi}\frac{\alpha_i P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)P(Y_{vi}=0|\theta_v;\boldsymbol{\iota}_i)}{P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)}+(1-y_{vi})\frac{-\alpha_i P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)P(Y_{vi}=0|\theta_v;\boldsymbol{\iota}_i)}{P(Y_{vi}=0|\theta_v;\boldsymbol{\iota}_i)}\right]\\
&= \sum_{i=1}^{k}\left[y_{vi}\alpha_i P(Y_{vi}=0|\theta_v;\boldsymbol{\iota}_i)-(1-y_{vi})\alpha_i P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)\right]\\
&= \sum_{i=1}^{k}\alpha_i\left[y_{vi}-P(Y_{vi}=1|\theta_v;\boldsymbol{\iota}_i)\right]\end{aligned}$$

2b. Die erste partielle Ableitung der des natürlichen Logarithmus der Normalverteilung wird beispielsweise bei der MAP-Schätzung der individuellen Werte der Personen bezüglich der latenten Variablen benötigt. Sei die Personenvariablen θ a priori normalverteilt mit $\theta \sim N(\mu_\theta,\sigma_\theta)$, dann ist der natürliche Logarithmus dieser Normalverteilung:

$$\begin{aligned}\ln[f(\theta)] &= \ln\left\{\frac{1}{\sqrt{2\pi}\sigma_\theta}\exp\left[-\frac{(\theta-\mu_\theta)^2}{2\sigma_\theta^2}\right]\right\}\\
&= \ln\left(\frac{1}{\sqrt{2\pi}\sigma_\theta}\right)+\ln\left(\exp\left[-\frac{(\theta-\mu_\theta)^2}{2\sigma_\theta^2}\right]\right)\\
&= \ln\left(\frac{1}{\sqrt{2\pi}\sigma_\theta}\right)-\frac{(\theta-\mu_\theta)^2}{2\sigma_\theta^2}\end{aligned}$$

Die letzte Zeile ergibt sich, da die e-Funktion und der natürliche Logarithmus Umkehrfunktionen voneinander sind, sodass allgemein gilt $\ln(\exp(f(\theta))) = f(\theta)$. Die erste Ableitung der Gleichung erfolgt unter Verwendung der Summenregel:

$$\begin{aligned}\frac{d}{d\theta}\ln[f(\theta)] &= \frac{d}{d\theta}\ln\left(\frac{1}{\sqrt{2\pi}\sigma_\theta}\right)+\frac{d}{d\theta}-\frac{(\theta-\mu_\theta)^2}{2\sigma_\theta^2}\\
&= 0-\left(\frac{d}{d\theta}\frac{1}{2\sigma_\theta^2}\theta^2-\frac{d}{d\theta}\frac{\mu_\theta}{\sigma_\theta^2}\theta+\frac{d}{d\theta}\frac{\mu_\theta^2}{2\sigma_\theta^2}\right)\\
&= -\left(\frac{1}{\sigma_\theta^2}\theta-\frac{\mu_\theta}{\sigma_\theta^2}\right)\\
&= -\frac{\theta-\mu_\theta}{\sigma_\theta^2}\end{aligned}$$

In Zeile 2 der Gleichung wird die Ableitung des ersten Summanden gleich null, da es sich um die Ableitung einer Konstanten handelt. Bei der Ableitung des zweiten Summanden wurden hier die binomischen Formeln zum Auflösen des quadrierten Klammerausdrucks verwendet. Nachfolgend wurde wiederum die Summenregel für resultierende quadratische Funktion von θ verwendet.