6 Analytische Untersuchungen

In diesem Kapitel werden wir die analytischen Ableitungen, die zur Bestimmung der Quadranten der Jacobi-Matrix des Gauss-Newton-Verfahrens mit Overlapfunktional benötigt werden, angeben. Mit M bezeichnen wir im folgenden die Menge der rellen Matrizen und geben die Dimensionen nur an, wenn es der Übersichtlichkeit dient.

6.1 Ableitungen des Overlap-Anteils

Sei der Overlap-Anteil gegeben

$$F^O = \sqrt{\frac{2\sigma_{mess}\sigma_{sim}}{\sigma_{mess}^2 + \sigma_{sim}^2}} \exp(-\frac{(\mu_{mess} - \mu_{sim})^2}{2(\sigma_{mess}^2 + \sigma_{sim}^2)})$$

6.1.1 Ableitungen nach μ_{sim} und σ_{sim}

Man leitet schnell her:

$$\frac{\partial F^O}{\partial \mu_{sim}} = F^O \cdot \frac{\mu_{mess} - \mu_{sim}}{\sigma_{mess}^2 + \sigma_{sim}^2} \tag{39}$$

$$\frac{\partial F^O}{\partial \sigma_{sim}} = F^O \cdot (\frac{\sigma_{mess}^4 - \sigma_{sim}^4 + \sigma_{sim}^2(\mu_{mess} - \mu_{sim})^2}{\sigma_{sim}(\sigma_{mess}^2 + \sigma_{sim}^2)^2}) \tag{40}$$

6.1.2 Ableitungen nach p

Für die Ableitung des Overlap nach den Parametern gilt:

$$\frac{\partial F_i^O}{\partial p} = \frac{\partial F_i^O}{\partial \mu_{sim,i}} \cdot \frac{\partial \mu_{sim,i}}{\partial p} + \frac{\partial F_i^O}{\partial \sigma_{sim,i}} \cdot \frac{\partial \sigma_{sim,i}}{\partial p}$$

Der erste Summand ergibt sich (Indexkennzeichnung sim aus Lesbarkeitsgründen wenn eindeutig vernachlässigt) aus

$$\frac{\partial F_i^O}{\partial \mu_{sim,i}} = \begin{pmatrix} \frac{\partial F_1^O}{\partial \mu_1} & .. & \frac{\partial F_1^O}{\partial \mu_N} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \mu_1} & .. & \frac{\partial F_N^O}{\partial \mu_N} \end{pmatrix} \in M_{N \times N}$$

und

$$\frac{\partial \mu_{sim,i}}{\partial p} = \begin{pmatrix} \frac{\partial \mu_1}{\partial p_1} & .. & \frac{\partial \mu_1}{\partial p_n} \\ & & \\ \frac{\partial \mu_N}{\partial p_{noPar}} & & \frac{\partial \mu_N}{\partial p_n} \end{pmatrix} = \text{Jacobi-Matrix des Standardverfahrens } \frac{\partial F^R}{\partial p} \in M_{N \times n}$$

zu

$$\frac{\partial F_i^O}{\partial \mu_{sim,i}} \cdot \frac{\partial \mu_{sim,i}}{\partial p} = \begin{pmatrix} \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \mu_k} \frac{\partial \mu_k}{\partial p_1} & .. & \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \mu_k} \frac{\partial \mu_k}{\partial p_n} \\ .. & & .. \\ \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \mu_k} \frac{\partial \mu_k}{\partial p_1} & .. & \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \mu_k} \frac{\partial \mu_k}{\partial p_n} \end{pmatrix} \in M_{N \times n}$$

$$= \begin{pmatrix} \frac{\partial F_1^O}{\partial \mu_1} \frac{\partial \mu_1}{\partial p_1} & .. & \frac{\partial F_1^O}{\partial \mu_1} \frac{\partial \mu_1}{\partial p_n} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \mu_N} \frac{\partial \mu_N}{\partial p_1} & .. & \frac{\partial F_N^O}{\partial \mu_N} \frac{\partial \mu_N}{\partial p_n} \end{pmatrix}$$

da alle anderen Summanden identisch 0. Für den zweiten Summanden betrachtet man

$$\frac{\partial F_i^O}{\partial \sigma_{sim,i}} = \begin{pmatrix} \frac{\partial F_1^O}{\partial \sigma_1} & .. & \frac{\partial F_1^O}{\partial \sigma_N} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \sigma_1} & .. & \frac{\partial F_N^O}{\partial \sigma_N} \end{pmatrix} \in M_{N \times N}$$

$$\frac{\partial \sigma_{sim,i}}{\partial p} = \begin{pmatrix} \frac{\partial \sigma_1}{\partial p_1} & .. & \frac{\partial \sigma_1}{\partial p_n} \\ & & \\ \frac{\partial \sigma_N}{\partial p_n} & & \frac{\partial \sigma_N}{\partial p_n} \end{pmatrix} \quad \text{(Indexkennzeichnung sim vernachlässigt)}$$

und damit gilt

$$\frac{\partial F_i^O}{\partial \sigma_{sim,i}} \cdot \frac{\partial \sigma_{sim,i}}{\partial p} = \begin{pmatrix} \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial p_1} & .. & \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial p_n} \\ .. & & .. \\ \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial p_1} & .. & \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial p_n} \end{pmatrix}$$

$$= \begin{pmatrix} \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial p_1} & .. & \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial p_n} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial p_1} & .. & \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial p_n} \end{pmatrix}$$

da alle anderen Summanden identisch 0.

Insgesamt ergibt sich

$$\frac{\partial F_i^O}{\partial p} = \begin{pmatrix} \frac{\partial F_1^O}{\partial \mu_1} \frac{\partial \mu_1}{\partial p_1} + \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial p_1} & .. & \frac{\partial F_1^O}{\partial \mu_1} \frac{\partial \mu_1}{\partial p_{noPar}} + \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial p_{noPar}} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \mu_N} \frac{\partial \mu_N}{\partial p_1} + \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial p_1} & .. & \frac{\partial F_N^O}{\partial \mu_N} \frac{\partial \mu_N}{\partial p_{noPar}} + \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial p_{noPar}} \end{pmatrix}$$

6.1.3 Ableitungen nach σ_p

Für den Ausdruck

$$\frac{\partial F_i^O}{\partial \sigma_p} = \frac{\partial F_i^O}{\partial \sigma_{sim,i}} \cdot \frac{\partial \sigma_{sim,i}}{\partial \sigma_p}$$

betrachten wir genauer:

$$\frac{\partial F_i^O}{\partial \sigma_{sim,i}} = \begin{pmatrix} \frac{\partial F_1^O}{\partial \sigma_{sim,1}} & .. & \frac{\partial F_1^O}{\partial \sigma_{sim,N}} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \sigma_{sim,1}} & .. & \frac{\partial F_N^O}{\partial \sigma_{sim,N}} \end{pmatrix} \in M_{NxN} \ ,$$

$\sigma_{sim,i}$ bedeutet hier die Varianz des Modells in jedem Punkt i

$$\frac{\partial \sigma_{sim,i}}{\partial \sigma_p} = \begin{pmatrix} \frac{\partial \sigma_{sim,1}}{\partial \sigma_{p1}} & .. & \frac{\partial \sigma_{sim,1}}{\partial \sigma_{pn}} \\ .. & & .. \\ \frac{\partial \sigma_{sim,N}}{\partial \sigma_{p1}} & .. & \frac{\partial \sigma_{sim,N}}{\partial \sigma_{pn}} \end{pmatrix} \in M_{Nxn} \ ,$$

σ_p bedeutet hier Varianz der Parameter

$$\frac{\partial F_i^O}{\partial \sigma_{sim,i}} \cdot \frac{\partial \sigma_{sim,i}}{\partial \sigma_p} = \begin{pmatrix} \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial \sigma_{p_1}} & .. & \sum_{k=1}^{N} \frac{\partial F_1^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial \sigma_{p_n}} \\ .. & & .. \\ \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial \sigma_{p_1}} & .. & \sum_{k=1}^{N} \frac{\partial F_N^O}{\partial \sigma_k} \frac{\partial \sigma_k}{\partial \sigma_{p_n}} \end{pmatrix} \in M_{Nxn}$$

$$= \begin{pmatrix} \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial \sigma_{p_1}} & .. & \frac{\partial F_1^O}{\partial \sigma_1} \frac{\partial \sigma_1}{\partial \sigma_{p_n}} \\ .. & & .. \\ \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial \sigma_{p_1}} & .. & \frac{\partial F_N^O}{\partial \sigma_N} \frac{\partial \sigma_N}{\partial \sigma_{p_n}} \end{pmatrix} \text{(Indexkennzeichnung sim in der Matrix vernachlässigt)}$$

da alle anderen Summanden identisch 0.

6.2 Analytische Ableitung der σ_{sim}

Die Modellvarianz Σ_y ergibt sich durch lineare Propagation der Parametervarianzen. Sei $J(p)$ die Jacobi-Matrix des Standardverfahrens, ausgewertet im Parameter p. Es seien Parametervarianzen σ_p eingeführt, die Varianz-Covarianzmatrix habe sich ergeben zu $\Sigma_p(\sigma_p)$. Nach (20) betrachten wir also

$$\Sigma_y = J(p) \cdot \Sigma_p(\sigma_p) \cdot J(p)^T \in M_{N \times N}$$

Es gilt

$$\sigma_{sim,i}^2 = \sum_{k=1}^{n} J_{ik}^2 \sigma_{pk}^2, \qquad , \ \sigma_{sim,i}^2 \text{ Diagonalelemente von } \Sigma_y \tag{41}$$

und daraus leitet man ab

$$\frac{\partial \sigma_{sim,i}}{\partial \sigma_{p_j}} = \frac{1}{\sigma_{sim,i}} (J_{ij}^2 \sigma_{p_j})$$

da die Ableitung $\frac{\partial \sigma_{p_k}}{\partial \sigma_{p_j}}$ nur für $j = k$ einen Beitrag ungleich Null liefert.

Weiter gilt

$$\frac{\partial \sigma_{sim}}{\partial p} = \begin{pmatrix} \frac{\partial \sigma_{sim,1}}{\partial p_1} & .. & \frac{\partial \sigma_{sim,1}}{\partial p_n} \\ .. & & .. \\ \frac{\partial \sigma_{sim,N}}{\partial p_1} & .. & \frac{\partial \sigma_{sim,N}}{\partial p_n} \end{pmatrix} \in M_{Nxn}$$

und

$$\frac{\partial \sigma_{sim,i}}{\partial p_j} = \frac{1}{\sigma_{sim,i}} (\sum_{k=1}^{n} J_{ik} \cdot \sigma_{p_j}^2 \cdot \frac{\partial J_{ik}}{\partial p_j}) \tag{42}$$

$$\text{mit } \frac{\partial J_{ik}}{\partial p_j} \text{ als zweiter Ableitung (Hesse-Matrix)!}$$

6.3 Übertragung der Normalverteilung von den Parametern auf die Zustände

Die Diagonalelemente der Modellvarianzmatrix Σ_y werden folgendermaßen gebildet:

$$\begin{aligned} \Sigma_y &= J(p) \cdot \Sigma_p(\sigma_p) \cdot J(p)^T \\ J(p) &= \begin{pmatrix} \frac{\partial y_1}{\partial p_1} & .. & \frac{\partial y_1}{\partial p_n} \\ .. & & .. \\ \frac{\partial y_N}{\partial p_1} & .. & \frac{\partial y_N}{\partial p_n} \end{pmatrix} \\ & \quad \text{Ableitung der Zustände nach den Parametern} \end{aligned} \tag{43}$$

$$\begin{aligned} \Sigma_p(\sigma_p) &= \begin{pmatrix} \sigma_1^2 & .. & 0 \\ .. & & .. \\ 0 & .. & \sigma_n^2 \end{pmatrix} \\ \Sigma_p(\sigma_p) \cdot J(p)^T &= \begin{pmatrix} \sigma_1^2 & .. & 0 \\ .. & & .. \\ 0 & .. & \sigma_n^2 \end{pmatrix} \cdot \begin{pmatrix} \frac{\partial y_1}{\partial p_1} & .. & \frac{\partial y_N}{\partial p_1} \\ .. & & .. \\ \frac{\partial y_1}{\partial p_n} & .. & \frac{\partial y_N}{\partial p_n} \end{pmatrix} \\ &= \begin{pmatrix} \sigma_1^2 \cdot \frac{\partial y_1}{\partial p_1} & .. & \sigma_1^2 \cdot \frac{\partial y_N}{\partial p_1} \\ .. & & .. \\ \sigma_n^2 \cdot \frac{\partial y_1}{\partial p_n} & .. & \sigma_n^2 \cdot \frac{\partial y_N}{\partial p_n} \end{pmatrix} \\ J(p) \cdot \Sigma_p(\sigma_p) \cdot J(p)^T &= \end{aligned} \tag{44}$$

$$\begin{pmatrix} \sigma_1^2 \cdot \frac{\partial y_1}{\partial p_1}^2 + ..\sigma_n^2 \cdot \frac{\partial y_1}{\partial p_n}^2 & .. & * \\ .. & .. & .. \\ * & .. & \sigma_1^2 \cdot \frac{\partial y_n}{\partial p_1}^2 + ..\sigma_n^2 \cdot \frac{\partial y_n}{\partial p_n}^2 \end{pmatrix}$$

$$\text{Diagonalelemente also } \sigma_{sim,i}^2 = \sum_{k=1}^{n} a_{ik}^2 \sigma_{pk}^2, \; a_{ik} = \frac{\partial y_i}{\partial p_k}, i = 1, ..N \tag{45}$$

6.4 Hesse-Matrix

Zur Berechnung der zweiten Ableitung der ursprünglichen Funktion nach den ursprünglichen Parametern (Hesse-Matrix) muss der entsprechende Differenzenquotient auf Grundlage einer vorliegenden Jacobi-Matrix der Zustände gebildet werden.

$$J = \begin{pmatrix} \frac{\partial y_1}{\partial p_1} & .. & \frac{\partial y_1}{\partial p_n} \\ .. & & .. \\ \frac{\partial y_N}{\partial p_1} & .. & \frac{\partial y_N}{\partial p_n} \end{pmatrix}$$

für einen gegebenen Parametersatz. Berechne für alle i=1,..n

$$J_{(p_i)} = \begin{pmatrix} \frac{\partial y_1}{\partial p_1} & .. & \frac{\partial y_1}{\partial p_n} \\ .. & & .. \\ \frac{\partial y_N}{\partial p_1} & .. & \frac{\partial y_N}{\partial p_n} \end{pmatrix}$$

für denselben Parametersatz, aber gestörtem $p_i : \widehat{p_i} = p_i + \Delta_H \cdot p_i$.

Zur Belegung von Δ_H siehe (5.1.3) und beachte, dass Δ_H von Δ abweichen muss.

$$\frac{\partial J}{\partial p_i} = \frac{J - J_{(p_i)}}{\Delta_H \cdot p_i}$$

Dies bedeutet einen Aufwand von n Funktionsauswertungen zur Bestimmung eines J, und damit einen Aufwand von $n \cdot n$ Funktionsauswertungen zur Bestimmung der vollständigen Hesse-Matrix.

Um diesen Aufwand zu umgehen, der in Modellen mit vielen Zustandsgrößen und Parametern zu einem Zeitproblem führt, wurde bereits zu Beginn des Projekts ein Broyden-Verfahren implementiert, welches die Hesse-Matrix approximiert. Da die Modelle, die in diesem Projekt endgültig bearbeitet wurden, von niedriger Dimension sind, konnte auf die Nutzung des Broyden-Verfahren verzichtet werden; dieses wird daher in dieser Arbeit nicht weiter beschrieben. Eine Abschätzung der Ungenauigkeiten, die durch das Broyden-Verfahren eingeführt werden, steht ebenso noch aus.