# Inverse Probleme in der Geophysik
# Vorlesung (Vertretung K. Spitzer)
# TU Bergakademie Freiberg, SS 2020

## Teil 9: Einführung in nichtlineare Inversionsprobleme


Thomas Günther (LIAG Hannover)
(Thomas.Guenther@extern.tu-freiberg.de)


22. Juni 2020

# Heutige Veranstaltung

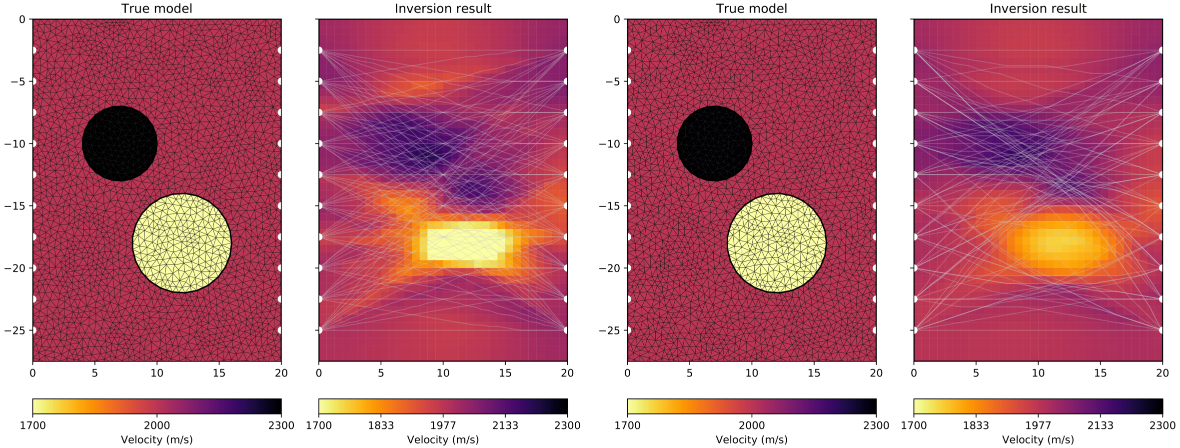
## Einführung nichtlineare Probleme

- Linearisierung und Lösungsverfahren (pdf)
  - Methode des Steilsten Abstiegs
  - Newton- und Gauss-Newton Methoden
- Modelltransformationen am Beispiel Strahlentomographie (pdf+jnb)
- Berechnung der Sensitivitätsmatrix (pdf)
- Beispielproblem 1D-Geoelektrik (jnb)

# Nichtlineare Inversionsprobleme

Bisher (linear): $\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}$ (linearer Zusammenhang zwischen Modell und Daten)
Meist hängen die Daten nicht-linear vom Modell ab!
Beispiel: Nichtlineare Strahlenwege hängen von Geschwindigkeit ab: $\mathbf{d} = \mathbf{G}(\mathbf{m})\mathbf{m} + \mathbf{n}$
Allgemeine Schreibweise: $\mathbf{d} = \mathbf{f}(\mathbf{m}) + \mathbf{n}$

# Nichtlineare Inversionsprobleme



# Nichtlineare Inversionsprobleme

Bisher (linear): $\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}$ (linearer Zusammenhang zwischen Modell und Daten)

Meist hängen die Daten nicht-linear vom Modell ab!

Beispiel: Nichtlineare Strahlenwege hängen von Geschwindigkeit ab: $\mathbf{d} = \mathbf{G}(\mathbf{m})\mathbf{m} + \mathbf{n}$

Allgemeine Schreibweise: $\mathbf{d} = \mathbf{f}(\mathbf{m}) + \mathbf{n}$

## Taylor-Entwicklung um Modell $\mathbf{m}^0$

$$\mathbf{f}(\mathbf{m}^0 + \Delta\mathbf{m}) \approx \mathbf{f}(\mathbf{m}^0) + \sum_j \frac{\partial f(\mathbf{m})}{\partial \mathbf{m}_j} \Delta\mathbf{m} = \mathbf{f}(\mathbf{m}^0) + \mathbf{S}\Delta\mathbf{m}$$

Die Jacobi-Matrix $\mathbf{S}$ enthält die Frechet-Ableitungen und wird auch Sensitivitätsmatrix genannt

Damit erhalten wir ein lineares Inversionsproblem (der Modell/Datendifferenzen)

$$\mathbf{S}\Delta\mathbf{m} = \Delta\mathbf{d} = \mathbf{d} - \mathbf{f}(\mathbf{m}^0)$$

Beachte: Taylor-Entwicklung gilt nur in der Nähe von $\mathbf{m}^0$! Also neu berechnen.

# Nichtlineare Inversionsprobleme

## Iterative Lösung nichtlinearer inverser Probleme

ausgehend von einem Start-Modell $\mathbf{m}^0$ errechnet man iterativ ($n$=0,1,2,...) immer neue Modelle

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \quad \Delta\mathbf{m}^n$$

durch Lösung des linearen Inversionsproblems

$$\mathbf{S}^n \Delta\mathbf{m}^n = \Delta\mathbf{d}^n = \mathbf{d} - \mathbf{f}(\mathbf{m}^n)$$

Alle weiteren Einflüsse (z.B. Regularisierung) können auf $\Delta\mathbf{m}$ angewandt werden.

# Nichtlineare Inversionsprobleme

## Iterative Lösung nichtlinearer inverser Probleme

ausgehend von einem Start-Modell $\mathbf{m}^0$ errechnet man iterativ ($n$=0,1,2,...) immer neue Modelle

$$\mathbf{m}^{n+1} = \mathbf{m}^n + \tau^n \Delta\mathbf{m}^n$$

durch Lösung des linearen Inversionsproblems

$$\mathbf{S}^n \Delta\mathbf{m}^n = \Delta\mathbf{d}^n = \mathbf{d} - \mathbf{f}(\mathbf{m}^n)$$

Alle weiteren Einflüsse (z.B. Regularisierung) können auf $\Delta\mathbf{m}$ angewandt werden.
Der Parameter $\tau$ wird auch als Schrittweite bezeichnet und kann optimiert werden, je nachdem wie weit die Taylor-Approximation gilt. Dieser Prozess wird als line search bezeichnet.

# Berechnung der Sensitivitätsmatrix

## Methoden

- analytische Berechnung (Ableitung von f)
- Lösung einer partiellen Differentialgleichung
- Perturbationsmethode (brute force)

$$\frac{\partial \mathbf{f}(\mathbf{m})}{\partial \mathbf{m}_j} \approx \frac{\mathbf{f}(\mathbf{m} + \delta_j \Delta m) - \mathbf{f}(\mathbf{m})}{\Delta m}$$

(jeden Modellparameter verändern und neue Zeile errechnen)

# Modelltransformationen

Sei $\mathbf{d} = \mathbf{G}\mathbf{m} + \mathbf{n}$ (z.B. Slowness)
Nehmen wir an, wir invertieren eine andere Größe $\hat{m}(m)$, z.B. Geschwindigkeit $v = 1/s$ anstatt Slowness. Dann ist das Problem nicht mehr linear.
Die Frechet-Ableitungen berechnen sich mit:

$$\frac{\partial f_i(\mathbf{m})}{\partial \hat{m}_j} = \frac{\partial f_i(\mathbf{m})}{\partial m_j} \frac{\partial m_j}{\partial \hat{m}_j}$$

Mit $\hat{m} = v = 1/s$ ergibt sich $\partial s/\partial v = -1/v^2 = -s^2$

## Jacobi-Matrix des transformierten Problems

$$\hat{S}_{ij} = \frac{\partial f_i(\mathbf{m})}{\partial \hat{m}_j} = \frac{\partial f_i(\mathbf{m})}{\partial \mathbf{m}_j} \frac{\partial m}{\partial \hat{m}} = \frac{\partial f_i(\mathbf{m})}{\partial \mathbf{m}_j} \Big/ \frac{\partial \hat{m}}{\partial m}$$

# Modelltransformationen

## Logarithmische Parameter

Geschwindigkeit oder Slowness? Spez. Widerstand oder Leitfähigkeit?
Oft invertiert man Logarithmen, da so gesichert ist, dass die Parameter positiv bleiben:
Mit $\hat{m} = \log s$ ergibt sich $\partial s / \partial \log s = 1/(\partial \log s / \partial s) = s$, also eine Skalierung von **S** mit $s$
Update: $\log m^{n+1} = \log m^n + \Delta m \Rightarrow m^{n+1} = m^n \cdot e^{\Delta m}$

## Parametergrenzen

Die Wahl des Logarithmus bietet eine natürliche untere Grenze 0 für das Modell, daher ist es so beliebt (physikalische Plausibilität)
Eine alternative untere Grenze kann gewählt werden mit $\hat{m} = \log(m - m_L)$
Eine obere Grenze kann gewählt werden mit $\hat{m} = \log(m_U - m)$, oder eine Kombination aus beiden:

$$\hat{m} = \log(m - m_L) - \log(m_U - m) = \log \frac{m - m_L}{m_U - m}$$

# Methode des steilsten Gradienten

Die Zielfunktion im Sinne kleinster Quadrate ist

$$\Phi(\mathbf{m}) = (\mathbf{d} - \mathbf{f}(\mathbf{m}))^T(\mathbf{d} - \mathbf{f}(\mathbf{m}))$$

Der negative Gradient

$$g(\mathbf{m}) = -\nabla_m \Phi = -\nabla_m (\mathbf{d} - \mathbf{f}(\mathbf{m}))^T(\mathbf{d} - \mathbf{f}(\mathbf{m})) = 2\mathbf{S}^T(\mathbf{d} - \mathbf{f}(\mathbf{m}))$$

gibt die Richtung an, in welche die Zielfunktion abnimmt.
Das entspricht der rechten Seite in den Normalgleichungen für $\Delta\mathbf{m}/\Delta\mathbf{d}$.

# Newton-Methode

Taylor-Entwicklung 2. Ordnung der Zielfunktion $\Phi(\mathbf{m}) = (\mathbf{d} - \mathbf{f}(\mathbf{m}))^T(\mathbf{d} - \mathbf{f}(\mathbf{m}))$

$$\Phi(\mathbf{m} + \Delta\mathbf{m}) = \Phi(\mathbf{m}) + (\nabla\Phi)^T \Delta\mathbf{m} + \frac{1}{2}\Delta\mathbf{m}^T(\nabla^2\Phi)^T\Delta\mathbf{m}$$

mit der Hesse-Matrix:

$$(\nabla^2\Phi)_{ij} = \frac{\partial^2\Phi}{\partial m_i \partial m_j}$$

Null-Setzen der ersten Ableitung ergibt:

$$\nabla^2\Phi\Delta\mathbf{m} = -2\nabla\Phi$$

Aber: Die Hesse-Matrix ist meist schwer zu berechnen

$$\nabla\nabla^T\Phi = \nabla(2\mathbf{S}^T(\mathbf{f}(\mathbf{m}) - \mathbf{d})) = 2\mathbf{S}^T\mathbf{S} + 2(\nabla^T\mathbf{S}^T)(\mathbf{f}(\mathbf{m}) - \mathbf{d})$$

# Gauss-Newton-Methode

Näherung der Hesse-Matrix mit

$$\nabla\nabla^T\Phi \approx 2\mathbf{S}^T\mathbf{S}$$

und Vernachlässigung höherwertiger Ableitungen führt zu

$$\mathbf{S}^T\mathbf{S}\Delta\mathbf{m} = \mathbf{S}^T(\mathbf{d} - \mathbf{f}(\mathbf{m})) = \mathbf{S}^T\Delta\mathbf{d}$$

also der Least-Squares-Lösung des linearisierten Problems

## Regularisierung von m (global)

erweiterte Zielfunktion

$$\Phi(\mathbf{m}) = (\mathbf{d} - \mathbf{f}(\mathbf{m}))^T(\mathbf{d} - \mathbf{f}(\mathbf{m})) + \lambda^2(\mathbf{C}\mathbf{m})^T(\mathbf{C}\mathbf{m})$$

führt zu

$$(\mathbf{S}^T\mathbf{S} + \lambda^2\mathbf{C}^T\mathbf{C})\Delta\mathbf{m} = \mathbf{S}^T(\mathbf{d} - \mathbf{f}(\mathbf{m})) - \lambda^2\mathbf{C}^T\mathbf{C}\mathbf{m}$$