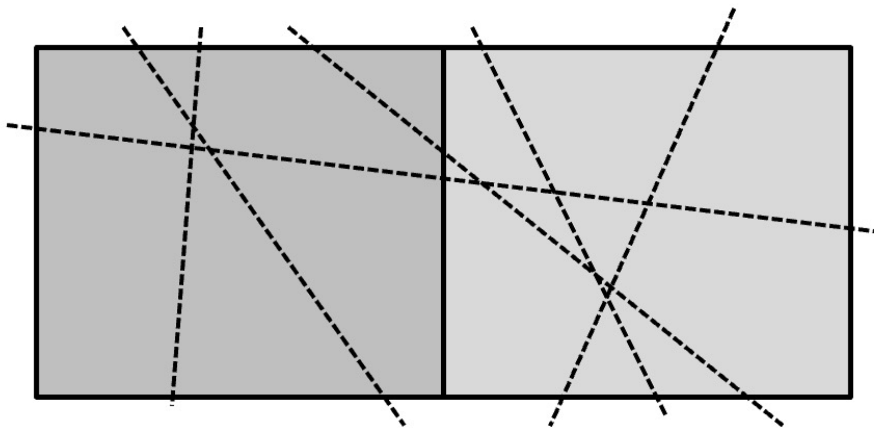
GEOL/PHYS 6670
Geophysical Inverse Theory

Lecture 4, September 14



Prof. Anne Sheehan

Anne.Sheehan @ colorado.edu

Office: Benson Rm 440A

Office Hours M 3:30-4:20, Tues 11-11:50

Homework 3 – due now

library exercise – look up another paper

coding practice – add and multiply matrices

Homework 4 – for next week

fitting a parabola with least squares

Class presentations

1) 10 minute presentation on a paper from the literature (the 2nd part of your HW 3). Starting next week. Sign up for a date here.

2) Term paper – handout on Canvas. Sign up for date and topic here.

GEOL/PHYS 6670 - Geophysical Inverse Theory

TERM PAPER ASSIGNMENT

Handout on Canvas

1. You may write a term paper on (a) an application of inverse theory to a problem of interest to you, including real computations or (b) on a technique in inverse theory that will not otherwise be covered in class, or (c) go into more depth on a topic that was covered in class. You will be required to give a 30-minute presentation in class on your term paper topic.

2. I have listed several acceptable topics for term papers below. Alternatively, you may select another related topic subject to my approval. Only one student per topic.
Sign up for topic and presentation date at
https://docs.google.com/spreadsheets/d/1meCF45qohFO7Ps5BOgbLd1IaZCJR_yBjMLak90ruH3M/edit?usp=sharing

3. Nominal length of reports is 5 double-spaced pages (roughly 2000 words) plus figures and bibliography (minimum of 4 references).

4. Due dates:
   - October 19 – Turn in short abstract describing topic chosen
   - November 16 - Student presentations of term projects begin
   - December 9 – Last peer reviews due (peer reviews due 2 days after each presentation)
   - December 11 - Term paper due

Topic List

The following is intended to give you an idea of some possible topics:

Last time –

Library, literature search info – Phil White
https://libguides.colorado.edu/geology
Menke Ch 2 (probability – variance and covariance)
Menke Ch 3 – L2 Norm and Least Squares

Today (lecture 4)

Continue Chapter 3: L2 norm and least squares

- weighted LS
- damped LS
- constrained LS

guiding principle for solving an inverse problem

find the $\mathbf{m}^{\text{est}}$

that minimizes $E = \|\mathbf{e}\|$

Called misfit, error, or cost function

with

$$\mathbf{e} = \mathbf{d}^{\text{obs}} - \mathbf{d}^{\text{pre}}$$

and

$$\mathbf{d}^{\text{pre}} = \mathbf{G}\mathbf{m}^{\text{est}}$$

$$E = \mathbf{e}^T\mathbf{e} = (\mathbf{d} - \mathbf{Gm})^T(\mathbf{d} - \mathbf{Gm}) = \sum_{i=1}^{N}\left[d_i - \sum_{j=1}^{M} G_{ij}m_j\right]\left[d_i - \sum_{k=1}^{M} G_{ik}m_k\right]$$

Taking derivative of E with respect to m and setting it equal to zero, eventually get

Least Square Solution

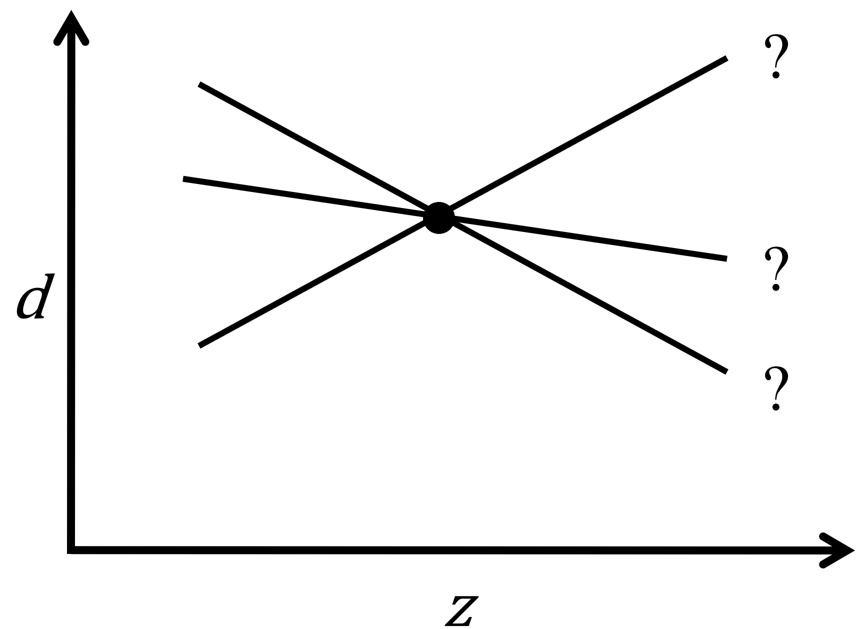
$$\mathbf{m}^{\mathrm{est}} = [\mathbf{G}^{\mathrm{T}}\mathbf{G}]^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{d}$$

$$\mathbf{m}^{\text{est}} = [\mathbf{G}^{\text{T}}\mathbf{G}]^{-1}\mathbf{G}^{\text{T}}\mathbf{d}$$

but Least Squares will fail when $[\mathbf{G}^{\text{T}}\mathbf{G}]$ has no inverse

When does it not have an inverse?
Consider straight line case…

example
fitting line to a single point



An infinity of different lines can pass through a single point.

For line fit, Gm –d is $\begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} d \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G}]^{-1} = \begin{bmatrix} N & \sum_{i=1}^{N} z_i \\ \sum_{i=1}^{N} z_i & \sum_{i=1}^{N} z_i^2 \end{bmatrix}^{-1} \rightarrow \begin{bmatrix} 1 & z_1 \\ z_1 & z_1^2 \end{bmatrix}^{-1}$$

Det (a, b; c, d) = ad – bc
$= z_1^2 - z_1^2 = 0$

zero determinant
hence no inverse

For line fit, Gm –d is $\begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} d \\ d_2 \\ \vdots \\ d_N \end{bmatrix}$

$$[\mathbf{G}^{\mathrm{T}}\mathbf{G}]^{-1} = \begin{bmatrix} N & \sum_{i=1}^{N} z_i \\ \sum_{i=1}^{N} z_i & \sum_{i=1}^{N} z_i^2 \end{bmatrix}^{-1} \rightarrow \begin{bmatrix} 1 & z_1 \\ z_1 & z_1^2 \end{bmatrix}^{-1}$$

zero determinant
hence no inverse

Det (a, b; c, d) = ad – bc
= $z_1^2 - z_1^2 = 0$

The straight line case fails when there is only one data point. The determinant of G^TG is zero in this case, and a matrix with zero determinant has no inverse.

Least Squares will fail

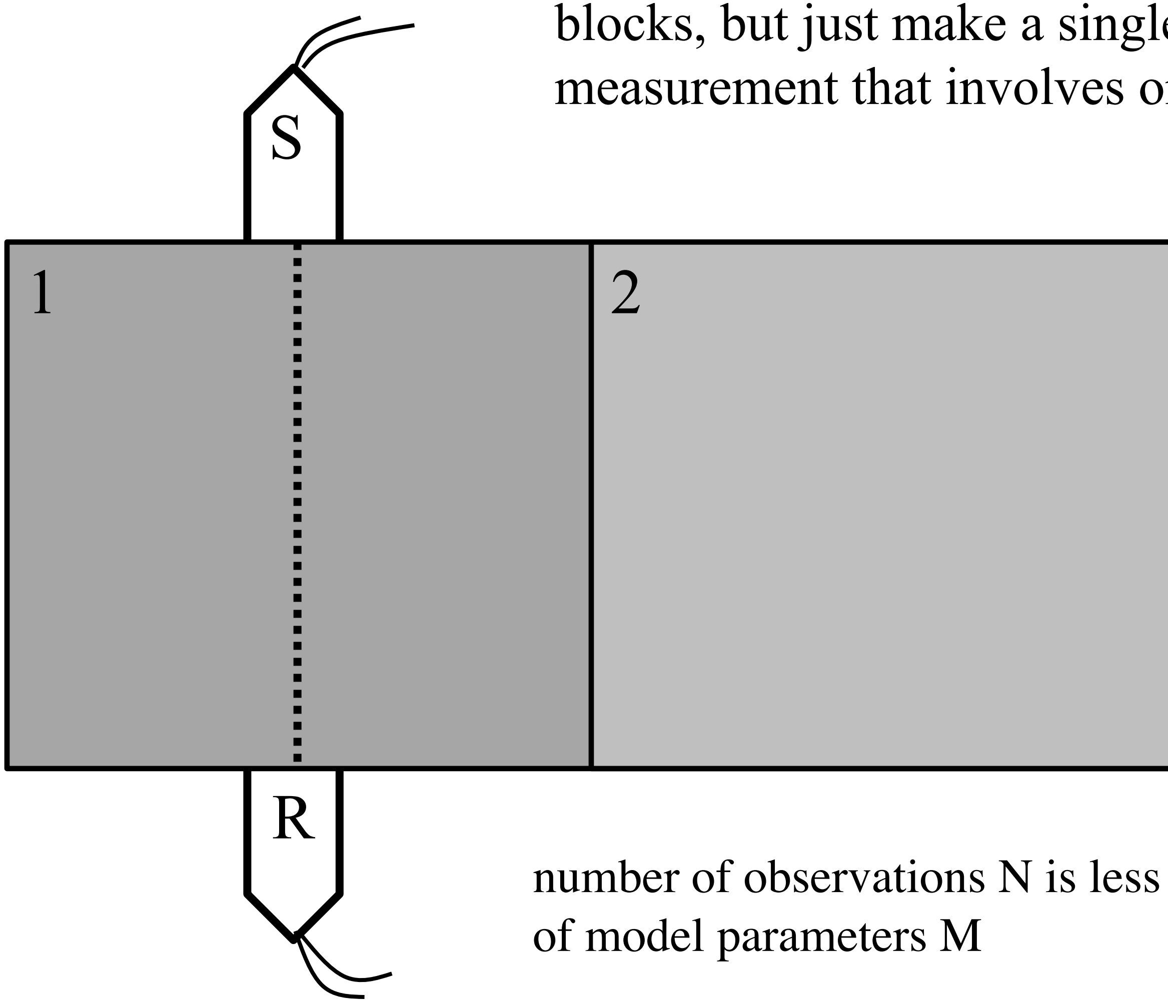
when more than one solution minimizes the error

the inverse problem is "underdetermined"

Underdetermined: not enough information to determine a unique solution.

example of an underdetermined problem

Want to know the properties of two blocks, but just make a single measurement that involves only the first.



number of observations N is less than number of model parameters M

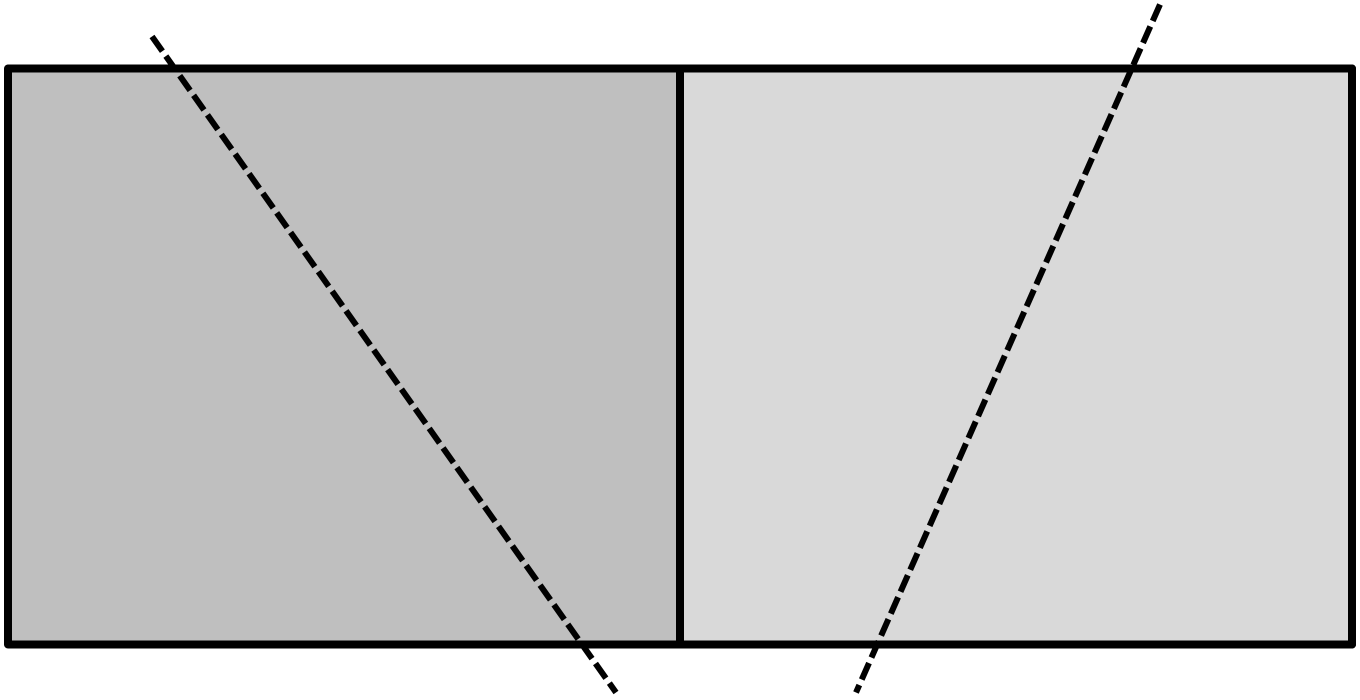
Underdetermined

Even determined

Overdetermined

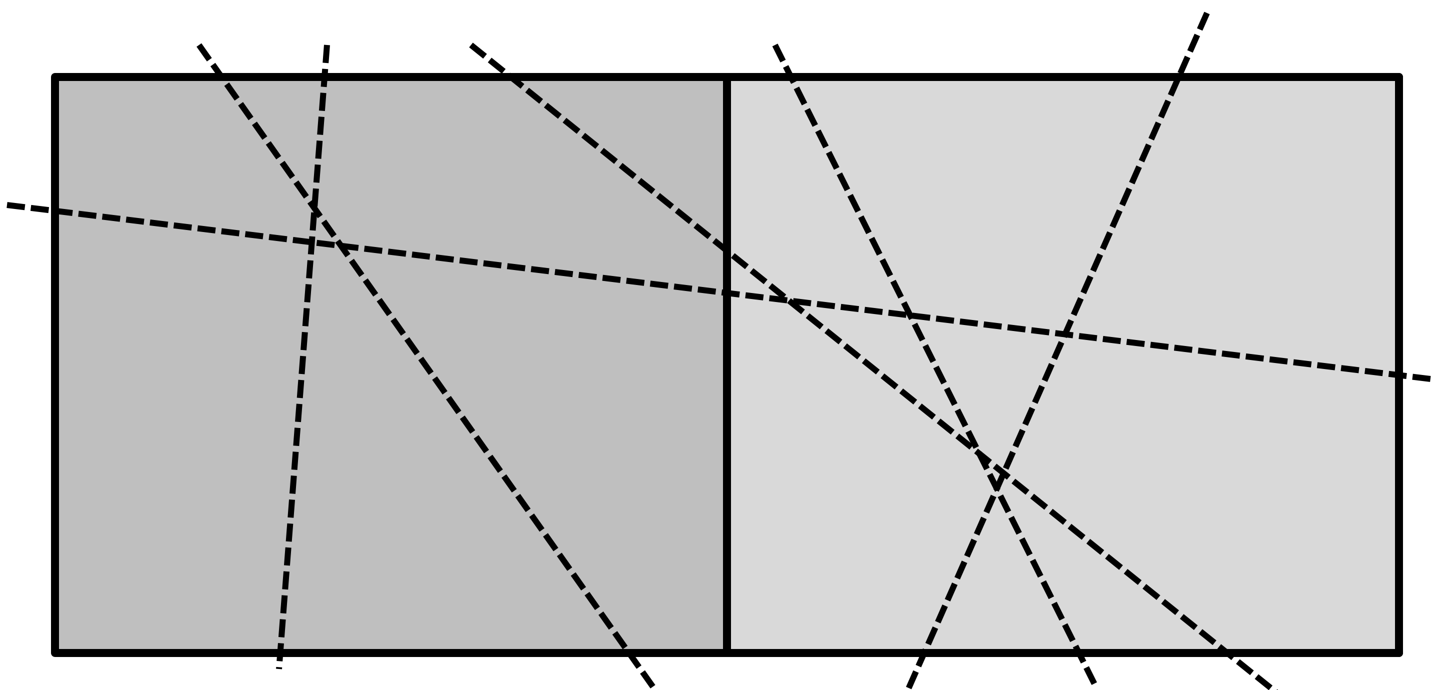
“Even Determined”

Exactly enough data is available to determine the model parameters



“Over Determined”

More than enough data is available to determine the model parameters



To solve **underdetermined** problems, we must add information that is not already in G.

This is called *a priori* information.
a priori, Latin for “from the former”
Assumptions based on prior knowledge

Examples of a priori information might include the constraint that density be greater than zero for rocks, or that the seismic P-wave velocity at the Moho falls within the range 5 - 10 km/s, etc.

More examples of a priori information

might choose to assume that model parameters are:

small
near a given value
have a known average value
smoothly varying with position
solve a known differential equation
positive
etc.

An example of a prior information: in some cases we might want to choose a solution that is small

$$\text{minimize } ||\mathbf{m}||_2$$

This is not the most sophisticated type of a priori information, but cases arise where it make sense (and other cases where it doesn't).

Find $\mathbf{m}^{est}$ that minimizes L=||$\mathbf{m}$||
subject to the constraint that $\mathbf{e}$= d – Gm =0

$$\Phi(\mathbf{m}) = L + \sum_{i=1}^{N} \lambda_i e_i = \sum_{i=1}^{M} m_i^2 + \sum_{i=1}^{N} \lambda_i \left[d_i - \sum_{j=1}^{M} G_{ij} m_j \right]$$

Minimize by taking the derivative of the 'cost function' and setting it equal to zero

$$\frac{\partial \Phi}{\partial m_q} = \sum_{i=1}^{M} 2\frac{\partial m_i}{\partial m_q} m_i - \sum_{i=1}^{N} \lambda_i \sum_{j=1}^{M} G_{ij} \frac{\partial m_j}{\partial m_q} = 2m_q - \sum_{i=1}^{N} \lambda_i G_{iq}$$

Where λ are Lagrange multipliers

In matrix form, $2\mathbf{m} = \mathbf{G}^{\mathrm{T}} \boldsymbol{\lambda}$ subject to $\mathbf{Gm} = \mathbf{d}$

Minimize by taking the derivative of the 'cost function' and setting it equal to zero

$$\frac{\partial \Phi}{\partial m_q} = \sum_{i=1}^{M} 2\frac{\partial m_i}{\partial m_q} m_i - \sum_{i=1}^{N} \lambda_i \sum_{j=1}^{M} G_{ij} \frac{\partial m_j}{\partial m_q} = 2m_q - \sum_{i=1}^{N} \lambda_i G_{iq}$$

Where λ are Lagrange multipliers

In matrix form, $2\mathbf{m} = \mathbf{G}^T \boldsymbol{\lambda}$ subject to $\mathbf{Gm} = \mathbf{d}$

algebraic manipulation of above yields

$$\tfrac{1}{2}\mathbf{G}\mathbf{G}^T \boldsymbol{\lambda} = \mathbf{d}$$

solve for λ

$$\boldsymbol{\lambda} = 2[\mathbf{G}\mathbf{G}^T]^{-1}\mathbf{d}$$

insert λ back into ($2\mathbf{m} = \mathbf{G}^T \boldsymbol{\lambda}$)

$$\mathbf{m} = \mathbf{G}^T [\mathbf{G}\mathbf{G}^T]^{-1}\mathbf{d}$$

Thus presuming $[\mathbf{G}\mathbf{G}^T]$ has an inverse

Minimum Length Solution

$$\mathbf{m}^{est} = \mathbf{G}^T [\mathbf{G}\mathbf{G}^T]^{-1} \mathbf{d}$$

minimizes $L=||\mathbf{m}||$ subject to the constraint that $\mathbf{e}= \mathrm{d} - \mathrm{Gm} =0$

Least Squares Solution

$$\mathbf{m}^{\text{est}} = [\mathbf{G}^{\text{T}}\mathbf{G}]^{-1}\mathbf{G}^{\text{T}}\mathbf{d}$$

Minimum Length Solution

$$\mathbf{m}^{\text{est}} = \mathbf{G}^{\text{T}}[\mathbf{G}\mathbf{G}^{\text{T}}]^{-1}\mathbf{d}$$

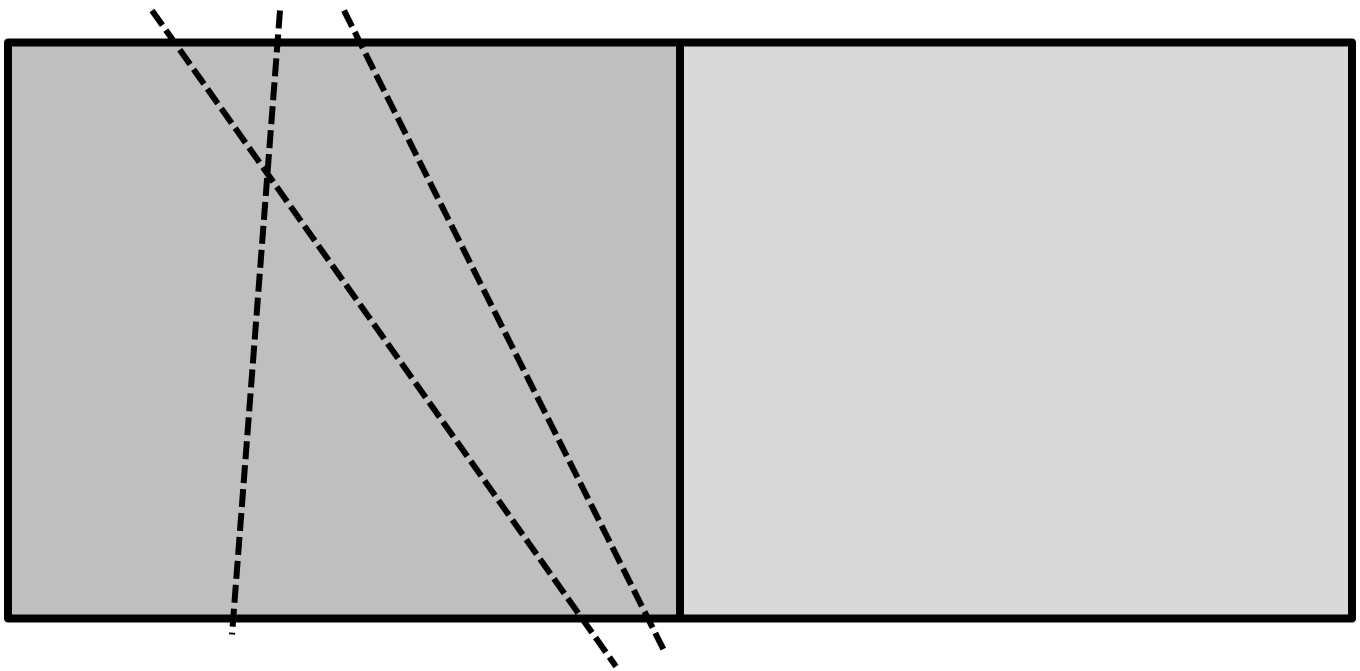
both have the linear form

$$\mathbf{m} = \mathbf{M}\mathbf{d}$$

“Mixed Determined”

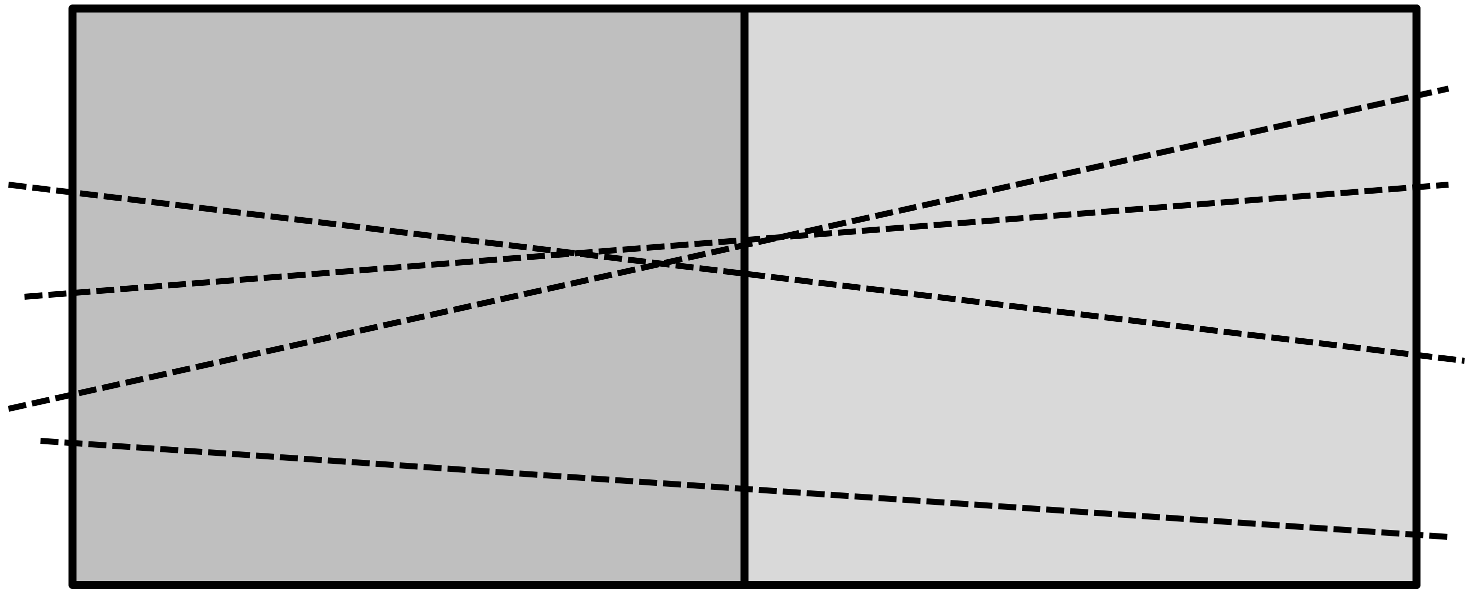
More than enough data is available to constrain some the model parameters

Insufficient data is available to constrain other model parameters



mixed-determined

the average of the two blocks is over-determined
and
the difference between the two blocks is under-determined



mixed-determined

some linear combinations of model parameters are not determined by the data

mixed-determined

some linear combinations of model parameters are not determined by the data

one approach to solving a mixed-determined problem

“of all the solutions that minimize $E=||\mathbf{e}||^2$ choose the one with minimum $L=||\mathbf{m}||^2$”

minimize

$$\Phi(\mathbf{m}) = E + \varepsilon^2 L = \mathbf{e}^{\mathrm{T}}\mathbf{e} + \varepsilon^2 \mathbf{m}^{\mathrm{T}}\mathbf{m}$$

“of all the solutions that minimize $E = ||\mathbf{e}||^2$ choose the one with minimum $L = ||\mathbf{m}||^2$”

minimize

$$\Phi(\mathbf{m}) = E + \varepsilon^2 L = \mathbf{e}^{\mathrm{T}}\mathbf{e} + \varepsilon^2 \mathbf{m}^{\mathrm{T}}\mathbf{m}$$

Leads to

damped least-squares solution

$$\mathbf{m}^{\mathrm{est}} = [\mathbf{G}^{\mathrm{T}}\mathbf{G} + \varepsilon^2 \mathbf{I}]\mathbf{G}^{\mathrm{T}}\mathbf{d}$$

Very similar to least-squares

$$\mathbf{m}^{\mathrm{est}} = [\mathbf{G}^{\mathrm{T}}\mathbf{G}]^{-1}\mathbf{G}^{\mathrm{T}}\mathbf{d}$$

Just add ε^2 to diagonal of $\mathbf{G}^{\mathrm{T}}\mathbf{G}$

Suppose that some data are more accurately determined than others

minimize

$$E = \mathbf{e}^{\mathrm{T}} \mathbf{W}_e \mathbf{e}$$

Where W_e is a weight matrix

example
when d_3 is more accurately measured than the other data

$$\mathbf{W}_e = \begin{bmatrix} 1 & & & & \\ & 1 & & & \\ & & 2 & & \\ & & & 1 & \\ & & & & 1 \end{bmatrix}$$

Weighted least squares
minimize E where

$$E = \mathbf{e}^{\mathrm{T}} \mathbf{W}_e \, \mathbf{e}$$

Weighted least squares solution

$$\mathbf{m}_{\mathrm{WLS}} = [\mathbf{G}^{\mathrm{T}} \mathbf{W}_e \mathbf{G}]^{-1} \mathbf{G}^{\mathrm{T}} \mathbf{W}_e \mathbf{d}$$

$\mathbf{W}_e$ error weight matrix, can represent one data type being more accurate than another

For minimum length solution we assumed **m** is small so minimized

$$L = \mathbf{m}^{\mathrm{T}}\mathbf{m}$$

But perhaps instead **m** is close to <**m**>

in that case minimize

$$L = (\mathbf{m} - \langle \mathbf{m} \rangle)^{\mathrm{T}} (\mathbf{m} - \langle \mathbf{m} \rangle)$$

Or in some situations we might want to assume that **m** varies slowly with position (**m** is flat)

characterize steepness with first-difference

$$\mathbf{l} = \begin{bmatrix} -1 & 1 & & & \\ & -1 & 1 & & \\ & & \ddots & \ddots & \\ & & & -1 & 1 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix} = \mathbf{Dm}$$

approximation for dm/dx

Or **that m** varies smoothly with position (**m** is smooth)

characterize roughness with second-difference

$$\mathbf{l} = \begin{bmatrix} 1 & -2 & 1 & & & & \\ & 1 & -2 & 1 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & 1 & -2 & 1 & \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix} = \mathbf{Dm}$$

approximation for d^2m/dx^2

m varies slowly/smoothly with position

minimize

$$L = \mathbf{l}^{\mathrm{T}}\mathbf{l} = [\mathbf{Dm}]^{\mathrm{T}}[\mathbf{Dm}] = \mathbf{m}^{\mathrm{T}}\mathbf{D}^{\mathrm{T}}\mathbf{Dm} = \mathbf{m}^{\mathrm{T}}\mathbf{W}_{m}\mathbf{m}$$

with $\mathbf{W}_{\mathrm{m}} = \mathbf{D}^{\mathrm{T}}\mathbf{D}$

weighted damped least squares minimize $E+\varepsilon^2 L$ with

$$L = [\mathbf{m} - \langle \mathbf{m} \rangle]^{\mathrm{T}} \mathbf{W}_m [\mathbf{m} - \langle \mathbf{m} \rangle]$$

and

$$E = \mathbf{e}^{\mathrm{T}} \mathbf{W}_e \mathbf{e}$$

<m> a priori values of model parameters
Wm model parameter weight matrix, can represent derivatives
We error weight matrix, can represent one data type being more accurate than another
ε determines the relative weight gives to model error and deviation from the a priori values

weighted damped least squares

minimize $E + \varepsilon^2 L$

with

$$L = [\mathbf{m} - \langle \mathbf{m} \rangle]^{\mathrm{T}} \mathbf{W}_m [\mathbf{m} - \langle \mathbf{m} \rangle]$$

and

$$E = \mathbf{e}^{\mathrm{T}} \mathbf{W}_e \mathbf{e}$$

weighted damped least squares solution

$$[\mathbf{G}^{\mathrm{T}} \mathbf{W}_e \mathbf{G} + \varepsilon^2 \mathbf{W}_m] \, \mathbf{m}^{\mathrm{est}} = \mathbf{G}^{\mathrm{T}} \mathbf{W}_e \, \mathbf{d} + \varepsilon^2 \mathbf{W}_m \langle \mathbf{m} \rangle$$

Still to cover from Chapter 3

- Model covariance matrix
- Constrained least squares

Model Covariance

4c

(easier

$$[\mathrm{cov}\, d] = \frac{1}{K} \sum (d_i^{(k)} - \bar{d}_i)(d_j^{(k)} - \bar{d}_j)$$

$$[\mathrm{cov}\, m] = \frac{1}{K} \sum (m_j^{(k)} - \bar{m}_j)(m_i^{(k)} - \bar{m}_i)$$

$Gm = d \; ; \; m = G^{-g} d$

$$[\mathrm{cov}\, m] = \frac{1}{K} \sum \left(G^{-g} d_i^{(k)} - G^{-g} \bar{d}_i\right)\left(G^{-g} d_j^{k} - G^{-g} \bar{d}_j\right)$$

$$= \frac{1}{K} \sum G^{-g} (d_i^{(k)} - \bar{d}_i)\, G^{-g} (d_j^{(k)} - \bar{d}_j)$$

$$= G^{-g} G^{-g} \frac{1}{K} \sum (d_i^{(k)} - \bar{d}_i)(d_j^{(k)} - \bar{d}_j)$$

$$(\mathrm{cov}\, m) = (G^{-g})(\mathrm{cov}\, d)(G^{-g})^T$$

$[\mathrm{cov}\, m]$

4d

for LS $\quad m^{est} = \underbrace{(G^T G)^{-1} G^T}_{G^{-g}} d$

$$[\mathrm{cov}\, m] = \cancel{\left(\left(G^T G\right)^{-1} G\right)(\mathrm{cov}\, d)\left(\left(G^T G\right)^{-1} G^T\right)^T}$$

$$G^{-g}\ \mathrm{cov}\, d\, (G^{-g})^T$$

$$= \left(\left(G^T G\right)^{-1} G^T\right)(\mathrm{cov}\, d)\left(\left(G^T G\right)^{-1} G^T\right)^T \overset{\text{skip?}}{=} (\mathrm{cov}\, d)(G^T G)^{-1}$$

$$= \cancel{(\mathrm{cov}\, d)(G^T G)^{-1}}$$

for min length,

$$m^{est} = \underbrace{G^T (G G^T)^{-1}}_{G^{-g}} d$$

$$(\mathrm{cov}\, m) = \left(G^T (G G^T)^{-1}\right) \mathrm{cov}\, d \left(G^T (G G^T)^{-1}\right)^T$$

$$\overset{\text{skip?}}{=} (\mathrm{cov}\, d)\left(G^T (G G^T)^{-2} G\right)$$

Least Squares Solution

$$\mathbf{m}^{\text{est}} = [\mathbf{G}^{\text{T}}\mathbf{G}]^{-1}\mathbf{G}^{\text{T}}\mathbf{d}$$

Minimum Length Solution

$$\mathbf{m}^{\text{est}} = \mathbf{G}^{\text{T}}[\mathbf{G}\mathbf{G}^{\text{T}}]^{-1}\mathbf{d}$$

both have the linear form

$$\mathbf{m} = \mathbf{M}\mathbf{d}$$

$$\mathbf{m} = \mathbf{M}\mathbf{d}$$

and

$$[\mathrm{cov}\ \mathbf{m}] = \mathbf{M}\ [\mathrm{cov}\ \mathbf{d}]\ \mathbf{M}^{\mathrm{T}}$$

if data are uncorrelated with uniform variance $\sigma_d^{\,2}$

$$[\mathrm{cov}\ \mathbf{d}] = \sigma_d^{\,2}\mathbf{I}$$

Least Squares Solution

$$[\text{cov}\ \mathbf{m}] = [\mathbf{G}^T\mathbf{G}]^{-1}\mathbf{G}^T\sigma_d^2\,\mathbf{G}[\mathbf{G}^T\mathbf{G}]^{-1}$$

$$[\text{cov}\ \mathbf{m}] = \sigma_d^2\,[\mathbf{G}^T\mathbf{G}]^{-1}$$

Minimum Length Solution

$$[\text{cov}\ \mathbf{m}] = \mathbf{G}^T\,[\mathbf{G}\mathbf{G}^T]^{-1}\,\sigma_d^2\,[\mathbf{G}\mathbf{G}^T]^{-1}\mathbf{G}$$

$$[\text{cov}\ \mathbf{m}] = \sigma_d^2\,\mathbf{G}^T\,[\mathbf{G}\mathbf{G}^T]^{-2}\mathbf{G}$$

Example of model covariance matrix

- Model covariance matrix for earthquake hypocenter solution what can we learn from model covariance?

example, eq location problem

model params

$$\begin{pmatrix} x \\ y \\ z \\ \text{origin time} \end{pmatrix}$$

model covariance matrix (from Stein & Wysession) p 421

$$(\text{covm}) = \begin{pmatrix} 0.06 & 0.01 & 0.01 & 0.00 \\ 0.01 & 0.08 & -0.13 & 0.01 \\ 0.01 & -0.13 & 1.16 & -0.08 \\ 0.00 & 0.01 & -0.08 & 0.01 \end{pmatrix}$$

what can we learn?

σ_{zz}^2, depth estimate variance, is g.t. σ_{xx}^2 and σ_{yy}^2

σ_{zt}^2 is negative, indicating tradeoff btwn focal depth & OT

get similar arrivals if earlier (t smaller) but deeper (z larger)

σ_{xy}^2 cov btwn x & y location uncertainties is nonzero

can extract a 2x2 submatrix $\begin{pmatrix} \sigma_{xx}^2 & \sigma_{xy}^2 \\ \sigma_{yx}^2 & \sigma_{yy}^2 \end{pmatrix}$ and use that to get error ellipses.

maj axis of error ellipse is from eigenvalues, direction from eigenvectors.

Constrained inversion

constrained LS

Menke section 3.10

include constraint eqns in $Gm=d$ and give it more weight

$$Fm = h$$

example: require mass of model parameters to equal some value h_1

Menke equation 3.51

$$\underline{\underline{F}}\,\underline{m} = \frac{1}{M}\begin{bmatrix} 1 & 1 & 1 & \cdots & 1 \end{bmatrix}\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix} = [h_1] = \underline{h}$$

# of model params (use M not N)

different constraint, require that a particular model param equals a certain value

$$\underline{\underline{F}}\,\underline{m} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \end{bmatrix}\begin{bmatrix} m_1 \\ m_2 \\ \vdots \\ m_M \end{bmatrix} = [h_1] = \underline{h}$$

(can use weighting, or use Lagrange multiplier method)

minimize $e^T e$ w/ constraint ~~$\underline{\underline{F}}\,\underline{m} =$~~ $\underline{\underline{F}}\,\underline{m} - \underline{h} = 0$

$$\phi = \underset{1\times d}{(Gm-d)^T}\;\underset{d\times 1}{(Gm-d)} + 2\lambda^T(Fm-h)$$

$$\partial\phi/\partial m = -2G^T d + 2G^T G m + 2\lambda^T F \quad \text{and set} = 0$$

2b

$$\frac{d\Phi}{dm} = 0 = -\cancel{2}G^T d + \cancel{2}G^T G m + \cancel{2}\lambda^T F = 0$$

solve this along with constraint eqn $Fm - h = 0$

rewrite $G^T G m + \lambda^T F = G^T d$ $\quad Fm = h$

$$3.55 \quad \boxed{\begin{bmatrix} G^T G & F^T \\ F & 0 \end{bmatrix} \begin{bmatrix} m \\ \lambda \end{bmatrix} = \begin{bmatrix} G^T d \\ h \end{bmatrix}}$$

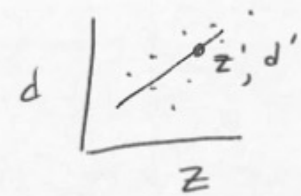
EXAMPLE

2c

3.10.1 constrained line fit
(p62)

$$d_i = m_1 + m_2 z_i$$

with constraint line must pass through z', d'



constraint eqn

$$d' = m_1 + m_2 z'$$

$$Fm = h$$

$$\text{(3.56)} \quad \underline{\underline{F}}\ \underline{m} = \begin{bmatrix} 1 & z' \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix} = \begin{bmatrix} d' \end{bmatrix} = \underline{h}$$

$$① \quad \begin{bmatrix} G^T G & F^T \\ F & 0 \end{bmatrix} \begin{bmatrix} m \\ \lambda \end{bmatrix} = \begin{bmatrix} G^T d \\ h \end{bmatrix}$$

~~use $G^T d$~~ : use $G^T G$ and $G^T d$ as we calculated a while back for the LS line fit example (p 46)

$$G \quad m = d$$

$$\begin{bmatrix} 1 & z_1 \\ 1 & z_2 \\ \vdots & \vdots \\ 1 & z_N \end{bmatrix}_{N \times 2} \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}_{2 \times 1} = \begin{bmatrix} d_1 \\ d_2 \\ \vdots \\ d_N \end{bmatrix}_{N \times 1}$$

2d

$$G^T G = \begin{bmatrix} N & \Sigma z_i \\ \Sigma z_i & \Sigma z_i^2 \end{bmatrix}$$

$$G^T d = \begin{bmatrix} \Sigma d_i \\ \Sigma z_i d_i \end{bmatrix}$$

plug into 1

$$\begin{bmatrix} N & \Sigma z_i & 1' \\ \Sigma z_i & \Sigma z_i^2 & z' \\ 1 & z' & 0 \end{bmatrix} \begin{bmatrix} m_1 \\ m_2 \\ \lambda \end{bmatrix} = \begin{bmatrix} \Sigma d_i \\ \Sigma z_i d_i \\ d' \end{bmatrix}$$

constrained line fit, w/ #5

2e

$y = a + bx$

constrain inversion to go through ~~point 3, x y~~ point 3, 4

$$x \quad y$$

$$y = a + bx$$

so $$4 = a + 3b$$

$$x \qquad y_2$$

$$F\,m = \begin{bmatrix} 1 & 3 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = h_1 = \cancel{3}\,4$$

$$\begin{bmatrix} G^T G & F^T \\ F & 0 \end{bmatrix} \begin{bmatrix} m \\ \lambda \end{bmatrix} = \begin{bmatrix} G^T d \\ h \end{bmatrix}$$

2x2 matrix

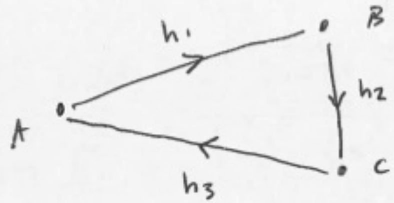
2x1 vector

$$\begin{bmatrix} \boxed{G^T G} & \begin{matrix} 1 \\ 3 \end{matrix} \\ \begin{matrix} 1 & 3 \end{matrix} & 0 \end{bmatrix} \begin{bmatrix} a \\ b \\ \lambda \end{bmatrix} = \begin{bmatrix} \begin{bmatrix} G^T d \end{bmatrix} \\ h \end{bmatrix}$$

3x3   3x1   3x1

another constrained LS example

consider a survey of 3 closed points A, B, & C measuring elevation diffs



path	length	Δ elevation
A B	18.1	25.42
B C	9.4	10.34
C A	14.2	−35.54
		1.22 total

should be ϕ but isn't!

Note elevation errors dn add to ϕ as they should.
Assume elevation errors are proportional to the path length.

Find elevation diffs h_1, h_2, h_3 st they add to zero (closure). do this by setting up LS problem w/a constraint eqn.
Must also acct for diff variances.

In this problem $Gm = d$ becomes

$$\underset{\underline{\underline{G}}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \underset{\underline{m}}{\begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix}} = \underset{\underline{d}}{\begin{bmatrix} d_1 \\ d_2 \\ d_3 \end{bmatrix}} = \begin{bmatrix} 25.42 \\ 10.34 \\ -35.54 \end{bmatrix}$$

use

$$W_e = \begin{bmatrix} \frac{1}{18.1} & 0 & 0 \\ 0 & \frac{1}{9.4} & 0 \\ 0 & 0 & \frac{1}{14.2} \end{bmatrix} = \begin{bmatrix} 0.06 & 0 & 0 \\ 0 & 0.11 & 0 \\ 0 & 0 & 0.07 \end{bmatrix}$$

The constraint is $Fm = h$

$$\begin{bmatrix} 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \end{bmatrix} = 0$$

Menke p. 61, eqn 3.55, but with weighting

$$\begin{bmatrix} G^T W_e G & F^T \\ F & 0 \end{bmatrix} \begin{bmatrix} m \\ \lambda \end{bmatrix} = \begin{bmatrix} G^T W_e d \\ h \end{bmatrix}$$

$$\begin{bmatrix} 0.06 & 0 & 0 & 1 \\ 0 & 0.11 & 0 & 1 \\ 0 & 0 & 0.07 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} h_1 \\ h_2 \\ h_3 \\ \lambda \end{bmatrix} = \begin{bmatrix} 0.06 \times 25.42 \\ 0.11 \times 10.34 \\ 0.07 \times (-35.54) \\ 0 \end{bmatrix} = \begin{bmatrix} 1.5252 \\ 1.1374 \\ -2.4878 \\ 0 \end{bmatrix}$$

extra

Solving a nonlinear inverse problem via Taylor Series expansion

(a way to linearize and solve for perturbations relative to an initial guess)

Solving a nonlinear inverse problem via Taylor Series expansion

Taylor series expansion
f(x) can be approximated by

$$f(x) =$$

$$f(a) + f'(a)(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f^{(3)}(a)}{3!}(x-a)^3 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + \ldots.$$

Taylor Series:

$f(m_0)$ +

$$d = f(m) = \left.\frac{df}{dm_1}\right| (m - m_0)$$

$$+ \left.\frac{df}{dm_2}\right| (m - m_0)$$

(partial derivatives)

$$+ \quad \frac{d^2 f}{dm_1^2}(m_1 - m_2)^2 + \cdots$$

$$= f(m_0) + \sum_{j=1}^{m} \frac{df(m)}{dm_j}(m - m_0)$$

$$(f(m_0) = initial\ guess)$$

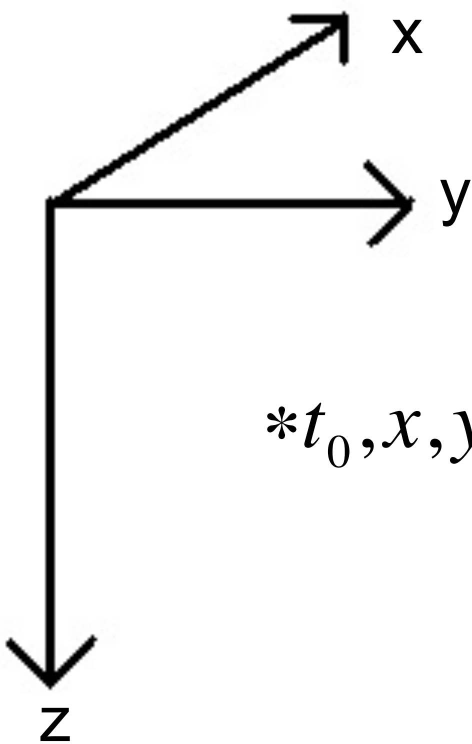
$$= f(m_0) + G_0(m - m_0)$$

$$f(m) - f(m_0) = G_0(m - m_0)$$

$$\Delta d = G\Delta m \quad iterate\ with\ new\ m_0$$

Nonlinear example 2:

EQ location



$(x_i, y_i) \rightarrow t_i$ *is travel time at* i^{th} *station*

$*t_0, x, y, z$

(1) Time @ ith station

Assuming station elevation is zero

$$t_i = t_0 + \frac{r_i}{\alpha}$$

$$= t_0 + \frac{\left[\left(x - x_i\right)^2 + \left(y - y_i\right)^2 + z^2\right]^{\frac{1}{2}}}{\alpha}$$

Want to solve for t_0, x, y, z

nonlinear

guess location $(\hat{x}, \hat{y}, \hat{z}, \hat{t}_0) = \ell_0$

this guess implies time t_i at station i

(2)

$$t_i = \hat{t}_0 + \left.\frac{d\hat{t}_i}{dt_0}\right|_{\ell_0} (t_0 - \hat{t}_0) + \left.\frac{d\hat{t}_i}{d_x}\right|_{\ell_0} (x - \hat{x}) + \left.\frac{d\hat{t}_i}{d_y}\right|_{\ell_0} (y - \hat{y}) + \left.\frac{d\hat{t}_i}{d_z}\right|_{\ell_0} (z - \hat{z})$$

$$t_i = t_0 + \frac{r_i}{\alpha}$$

$$= t_0 + \frac{\left[\left(x - x_i\right)^2 + \left(y - y_i\right)^2 + z^2\right]^{\frac{1}{2}}}{\alpha}$$

(3) Take derivative of equation (1)

$$\left.\frac{dt_i}{dt_0}\right|_{\ell_0} = 1$$

$$\left.\frac{dt_i}{dx}\right|_{\ell_0} = \frac{1}{\alpha} \left.\frac{\hat{x} - x_i}{r_i}\right|_{\ell_0} \qquad etc.$$

In form Gm = d,

$$\underbrace{\begin{pmatrix} 1 & \frac{1}{\alpha}\frac{\hat{x} - x_i}{r_i} & \frac{1}{\alpha}\frac{\hat{y} - y_i}{r_i} & \frac{1}{\alpha}\frac{\hat{z} - z_i}{r_i} \\ 1 & & & \\ 1 & & & \\ 1 & & & \end{pmatrix}}_{G} \underbrace{\begin{pmatrix} t_0 - \hat{t}_0 \\ x - \hat{x} \\ y - \hat{y} \\ z - \hat{z} \end{pmatrix}}_{m} = \underbrace{\begin{pmatrix} t_1 - \hat{t}_1 \\ t_2 - \hat{t}_2 \\ \\ t_n - \hat{t}_n \end{pmatrix}}_{d}$$

Steps for nonlinear iterative inversions:

1) make initial guess of model parameters, m_0
2) calculate partial derivatives evaluated at m_0, this gives the matrix G
3) Since

$$G\Delta m = d$$

solve for

$$\Delta m = (G^T G)^{-1} G^T \Delta d$$

4) solve

$$m_1 = m_0 + \Delta m$$

let this be the new model, go to step (1) and iterate

stop when residual $\sum \left(d_{pred} - d_{obs}\right)^2$

is small enough.