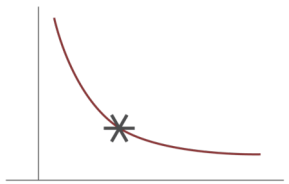
Linear Tikhonov Inversions

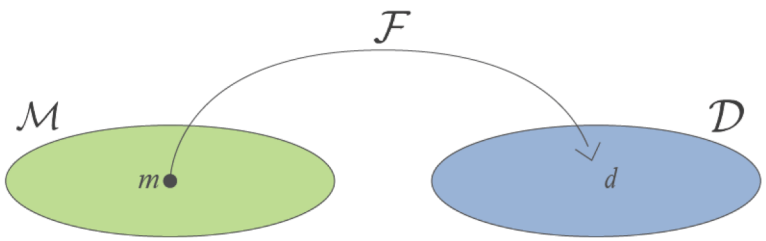


Outline

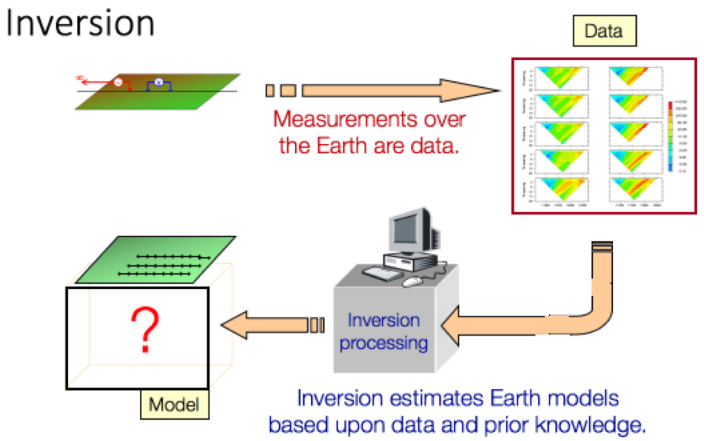
- Statement of the inverse problem
- App for a linear problem
- Data misfit
- Model norms and their importance
- Tikhonov approach to inversion
- A flow chart to keep you on track

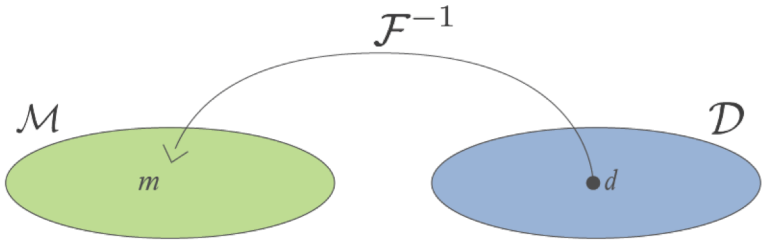
Our statement of the inverse problem

- Given observations: $d_j^{obs}, \quad j = 1, \ldots, N$
  - Uncertainties: ϵ_j
  - Ability for forward modelling: $\mathcal{F}[m] = d$



- Find the earth model that gave rise to the data.

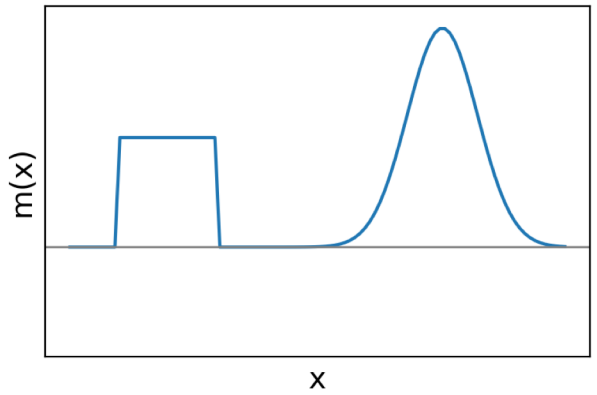


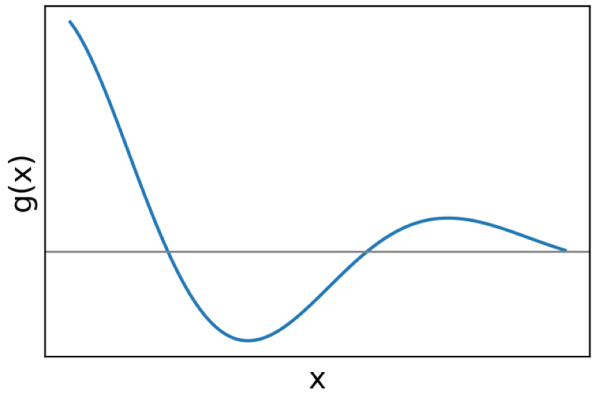


Linear problem

$$d_j = \int_v g_j(x) m(x) dx \qquad g_j(x) = e^{jpx} cos(2\pi jqx)$$

- d_j : j-th datum
- g_j : kernel function for j-th datum
- m : model





Each datum is an inner product of the model with the kernel function.

For this case

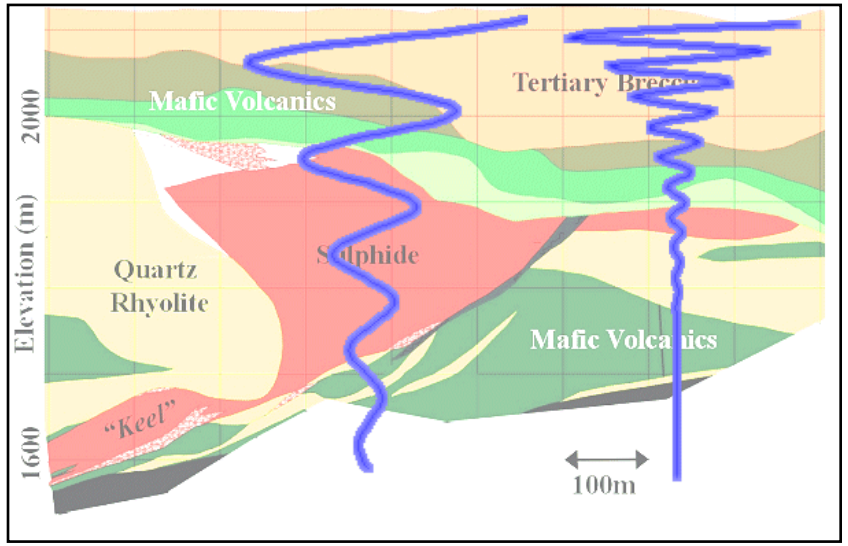
$$d_j = \mathbf{g} \cdot \mathbf{m} = 4.89$$

Analogy with 1D frequency domain EM

- FDEM system (Resolve)
- Signals: sinusoids at 5 frequencies
- Penetration depth depends upon frequency
- Measurements are fields from buried conductors

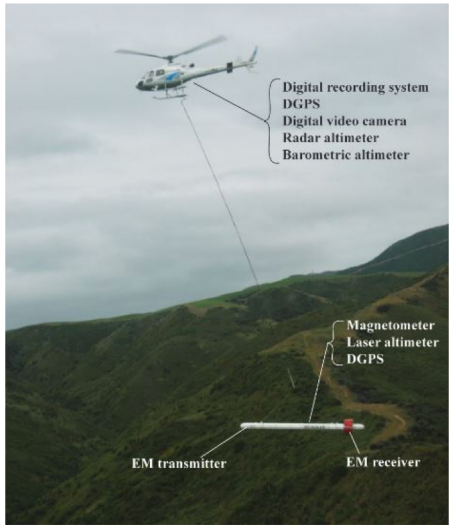
Low frequency

High frequency

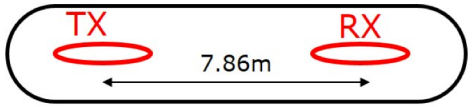


$$\delta = 503\sqrt{\frac{\rho}{f}}$$

Resolve system (2008)



Horizontal Co-planar



HCP frequencies:
382, 1822, 7970, 35920 and 130100 Hz

$$g_j(x) = e^{jpx}cos(2\pi jqx)$$

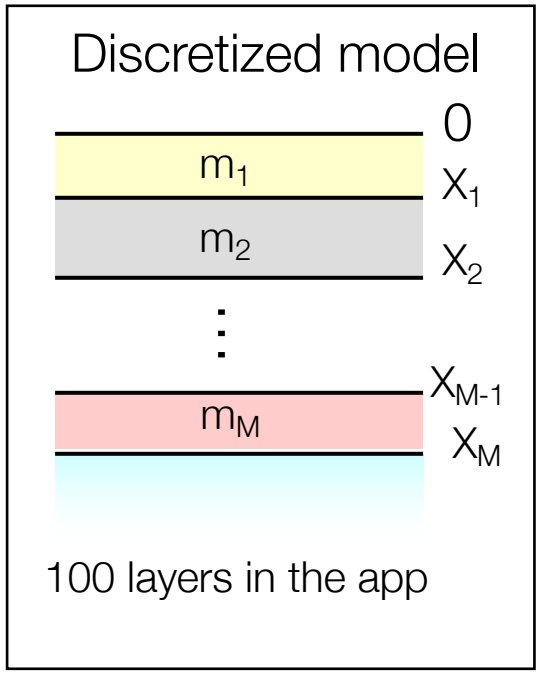
Forward Problem

- Datum is defined as: $d_i = \int_0^1 g_i(x) m(x) dx$
- Discretize model: $\mathbf{m} = (m_1, m_2, ..., m_M)$

$$d_i = \int_0^{X_1} g_i(x) m_1 dx + \int_{X_1}^{X_2} g_i(x) m_2 dx + ...$$

$$= \sum_{j=1}^{M} \Big(\int g_i(x) dx \Big) m_j$$

- In matrix form: $\mathbf{d} = \mathbf{G}\mathbf{m}$



$\mathbf{G}$: $(N \times M)$ matrix
$\mathbf{d}$: $(N \times 1)$ vector
$\mathbf{m}$: $(M \times 1)$ vector

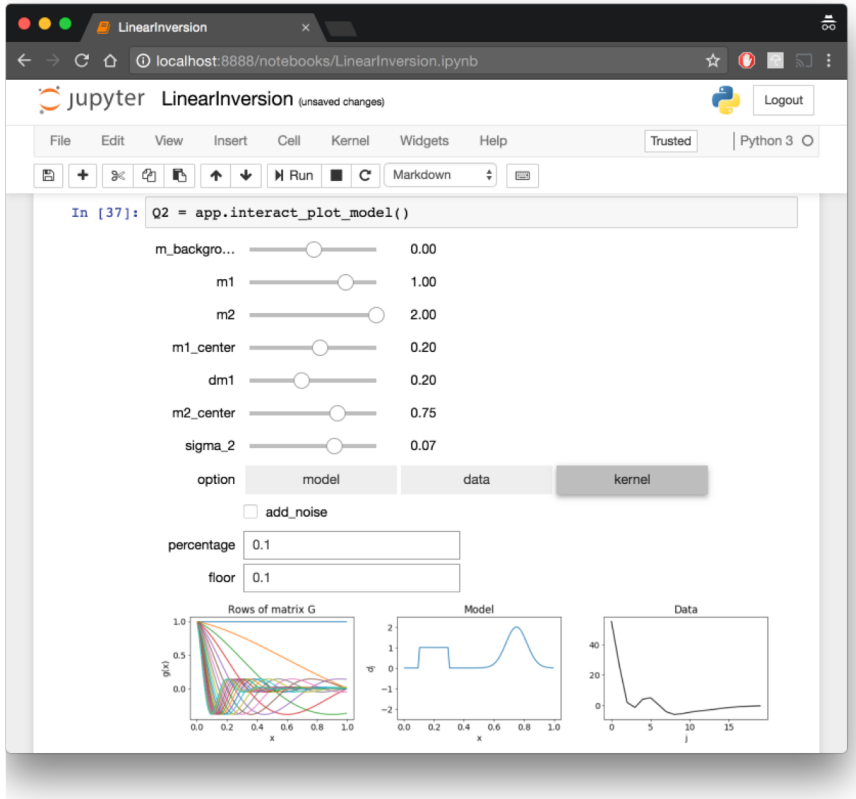
Linear inversion app (demo)

- It will help us
  - Develop a model
  - Consider kernels
  - Generate data
- Model: $m(x)$
- Kernels (physics):

$$g_j(x) = e^{jpx} cos(2\pi jqx)$$

- Data:

$$d_j = \int_v g_j(x) m(x) dx$$

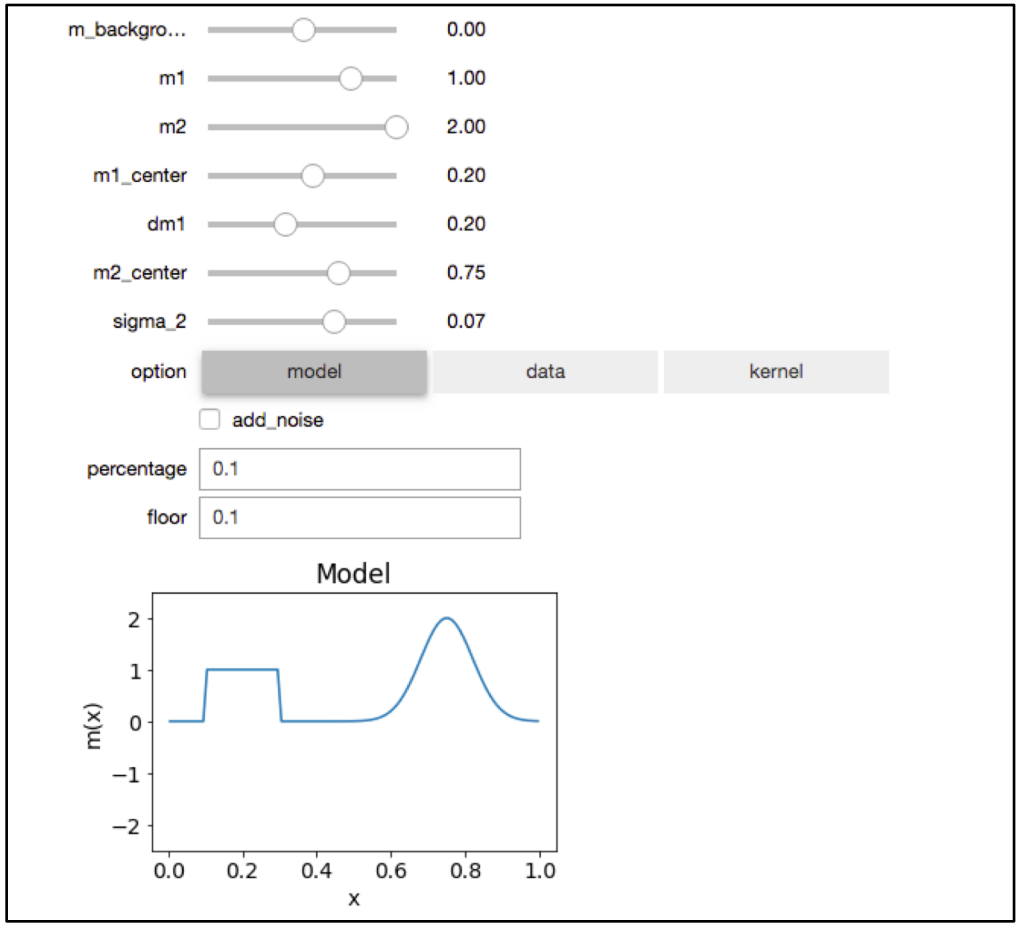


http://em.geosci.xyz/apps.html

Models with the app (demo)

Build a 1D model

- Background
- Box car
- Gaussian

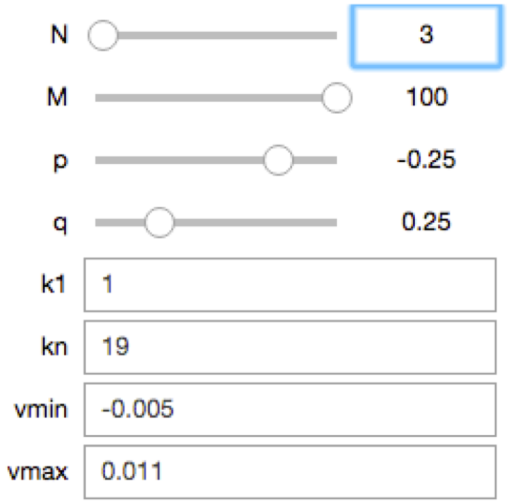


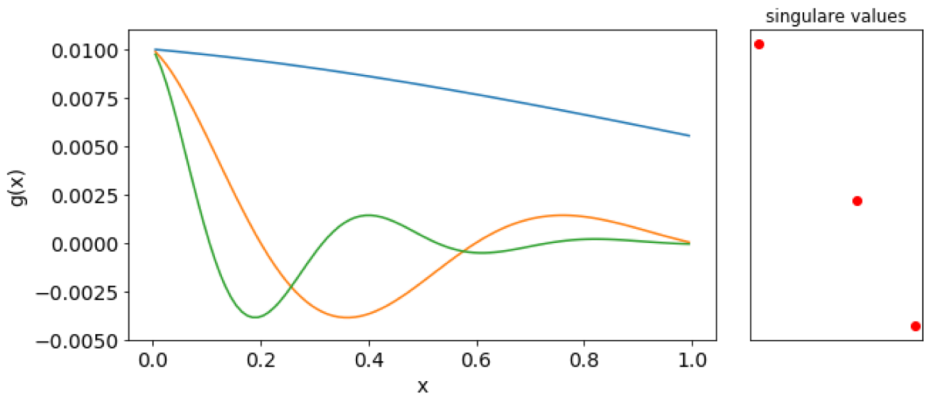
Kernels with the app (demo)

- Kernels for FEM are decaying sinusoids

$$g_j(x) = e^{jpx} cos(2\pi jqx)$$

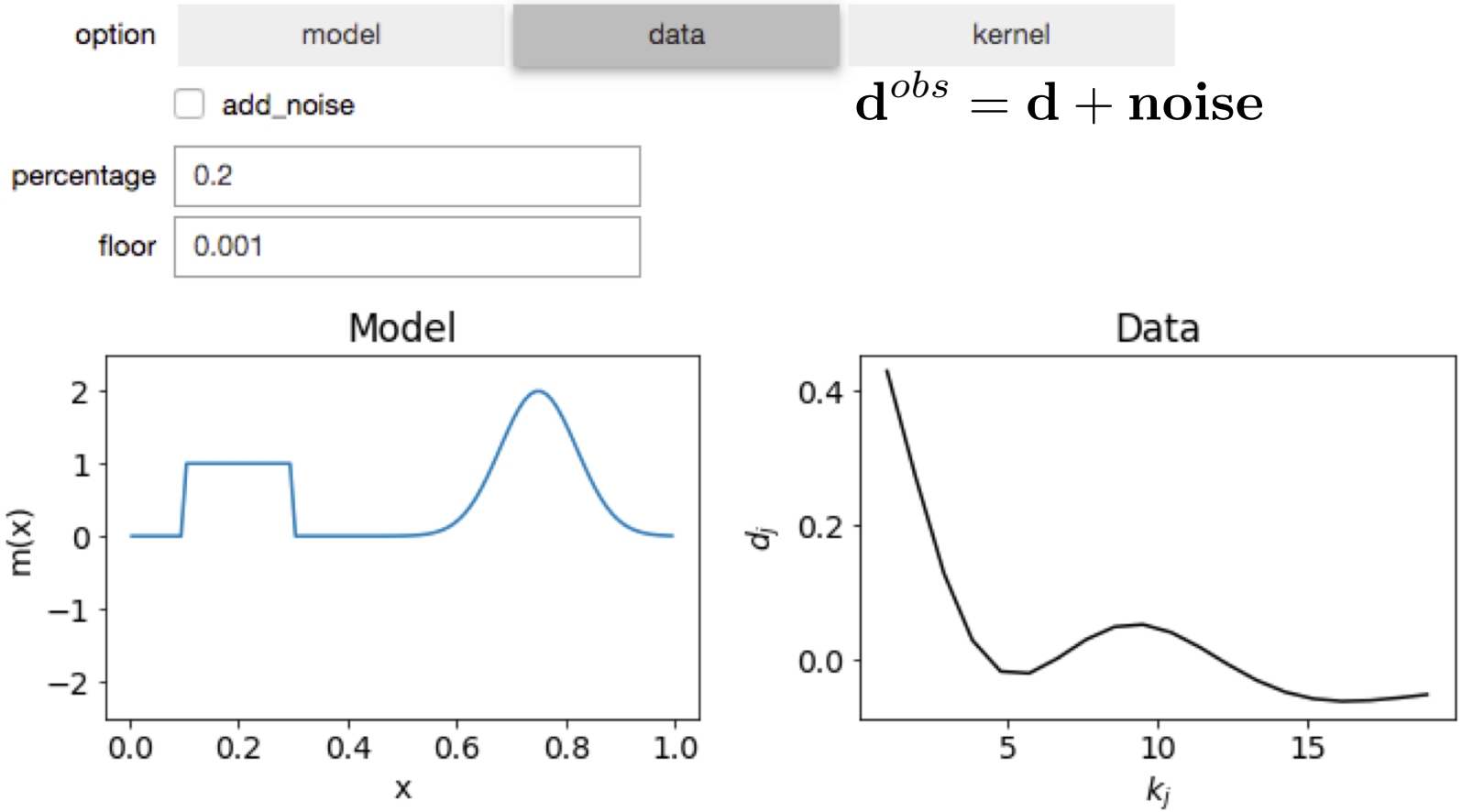
- N: number of data
- M: number of cells
- p: controls decay
- q: frequency
- j1: starting j value
- jN: ending j value
- ymin, ymax (plot limits)





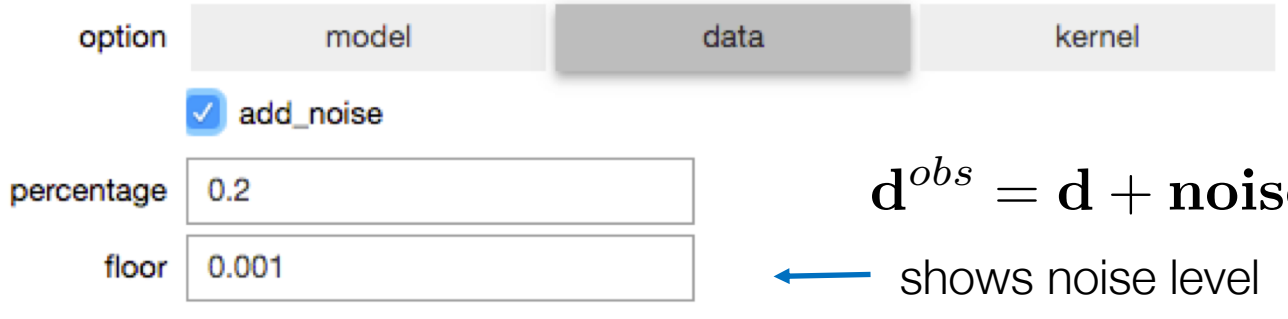
Real observation includes noise

- Data: $\mathbf{d} = \mathbf{Gm}$



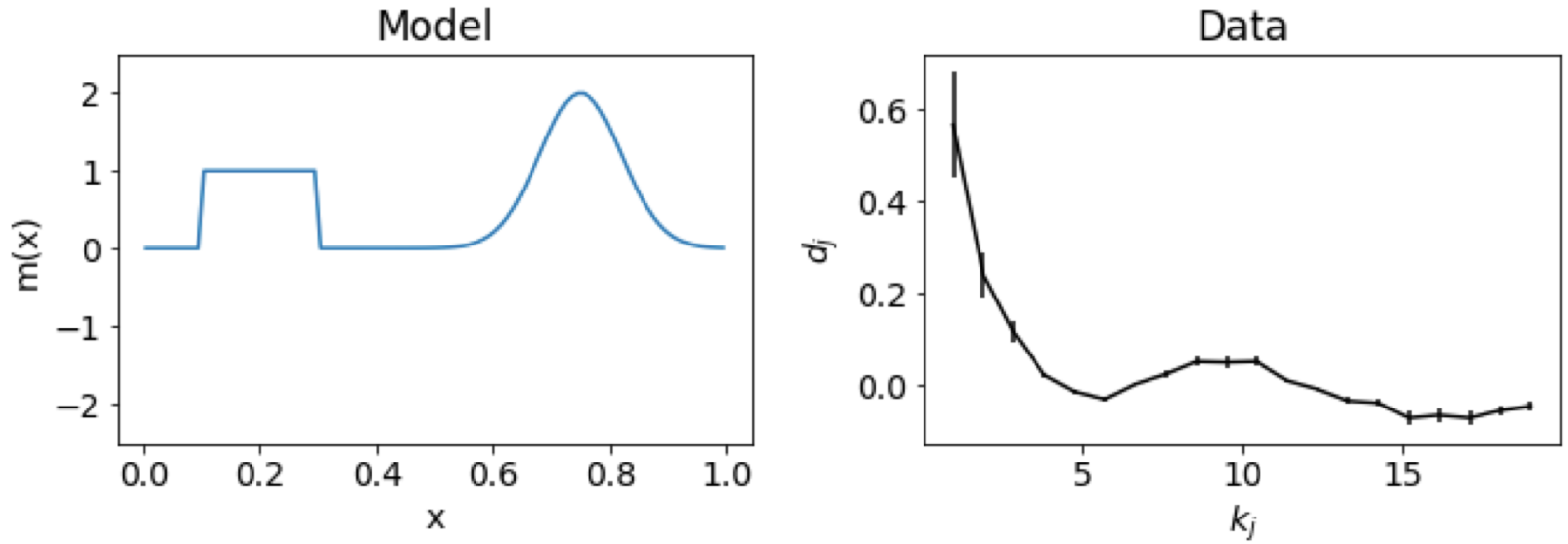
Real observation includes noise

- Data: $\mathbf{d} = \mathbf{Gm}$

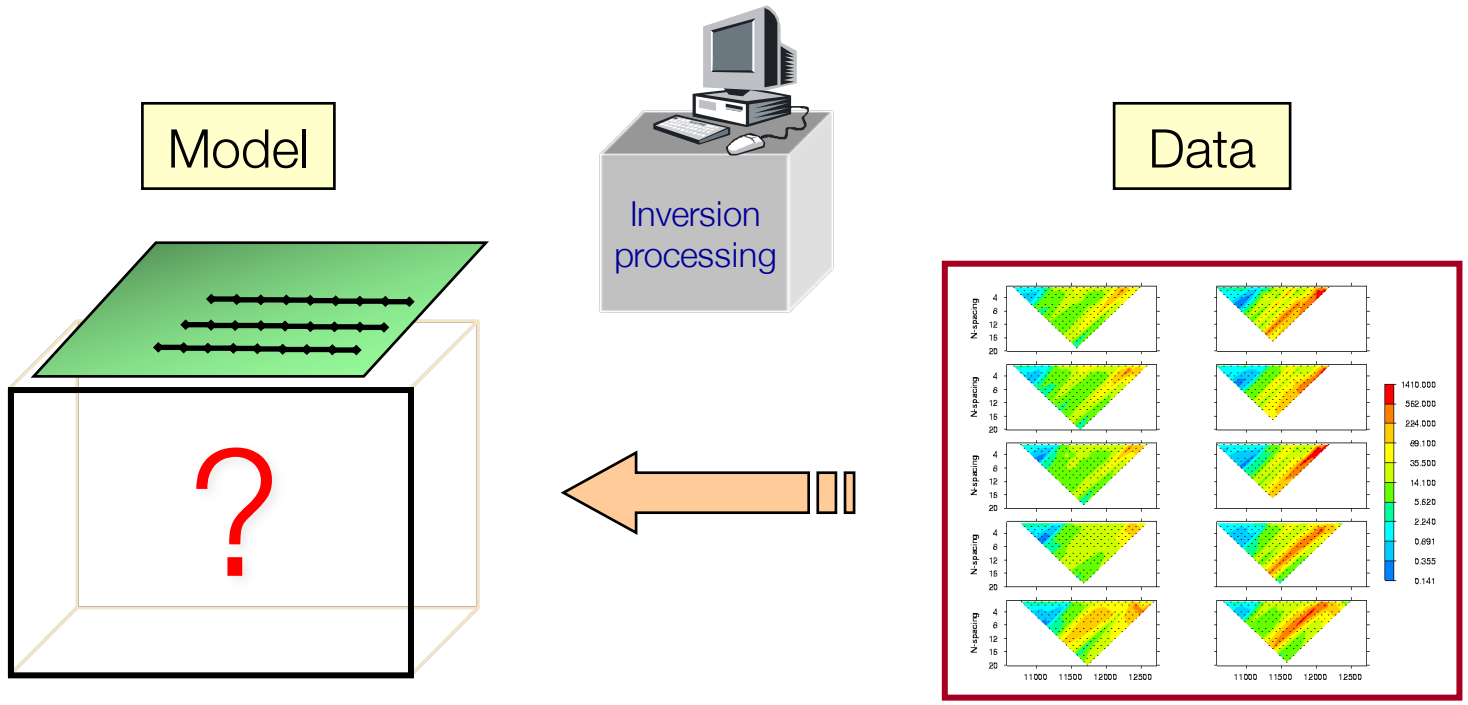


$$\mathbf{d}^{obs} = \mathbf{d} + \mathbf{noise}$$

← shows noise level

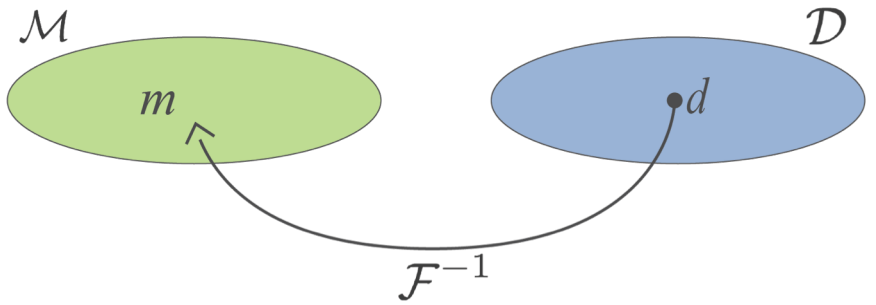


Inversion



Inverse problem

- Observed data: $d_j^{obs}, \quad j = 1, \ldots, N$
- Uncertainty: ϵ_j
- Ability to simulate data: $\mathcal{F}[m] = d$
- Find the model which fits the observation



- For linear problem: $\mathbf{Gm} = \mathbf{d}$

$\mathbf{m} \in \mathbb{R}^M$
$\mathbf{d} \in \mathbb{R}^N$
$\mathbf{G} \in \mathbb{R}^{M \times N}$

$M > N$

more unknowns than data
(underdetermined system)

Inversion using Misfit criterion

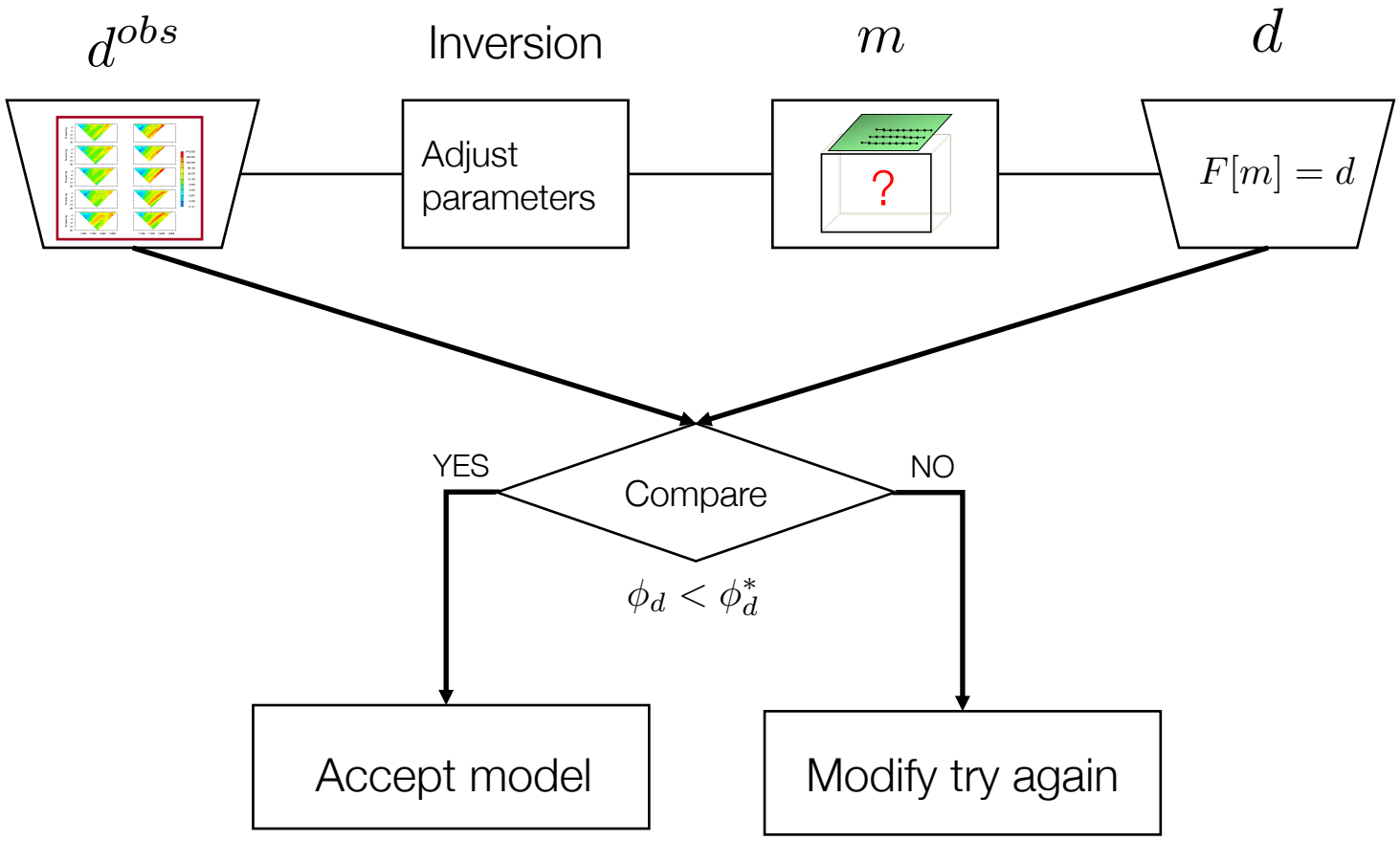
Forward modelling	$\mathbf{Gm} = \mathbf{d}$
Data	$\mathbf{d}^{obs} = \mathbf{d} + \boldsymbol{\epsilon}$
Noise	$\boldsymbol{\epsilon}$

- Gaussian errors with standard deviation, ϵ_j
- Misfit measure: $\phi_d = \sum_{j=1}^{N} \left(\frac{d_j - d_j^{obs}}{\epsilon_j} \right)^2$
- Expected value of ϕ_d is $E[\phi_d] = N$
- Data are fit when $\phi_d \simeq \phi_d^*$

ϕ_d^*: target misfit

$\phi_d^* = N$

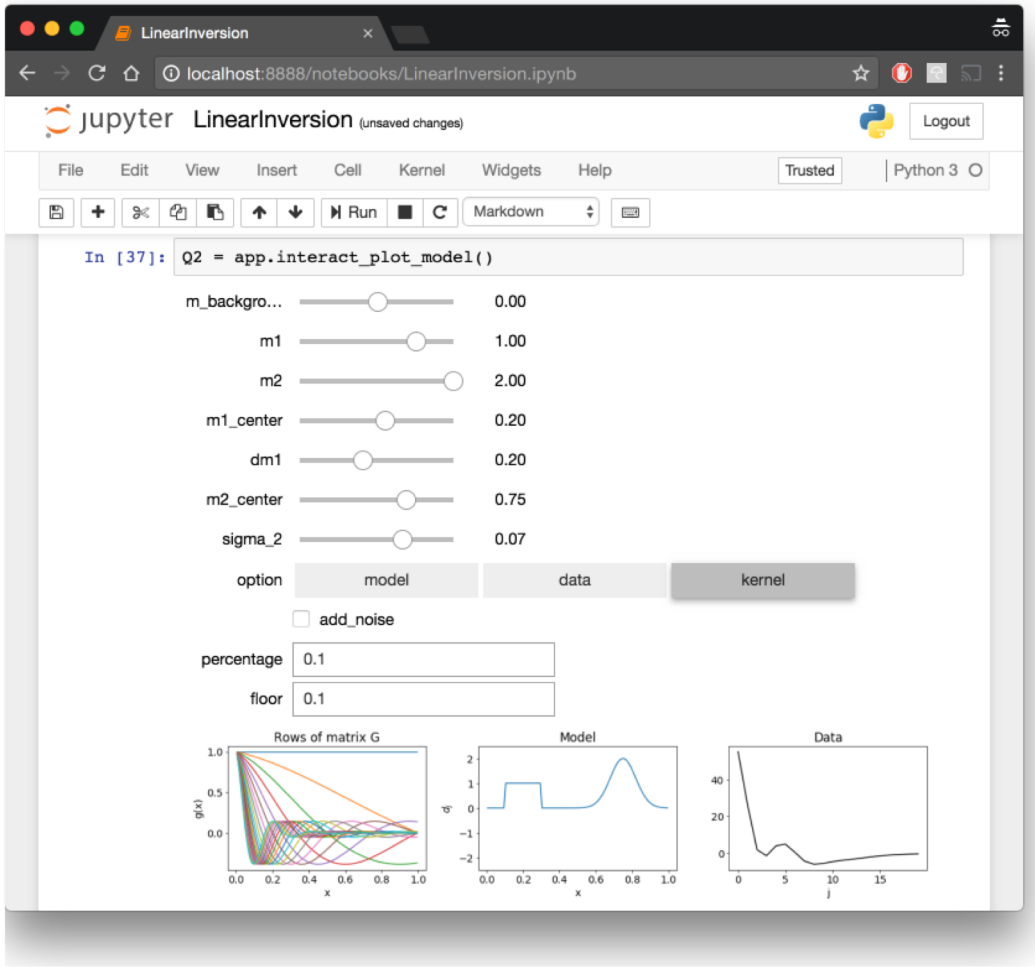
Inversion with misfit only



$$\phi_d = \sum_{j=1}^{N} \left(\frac{d_j - d_j^{obs}}{\epsilon_j} \right)^2$$

Inversion app (demo)

- With accurate data
- With noisy data.



The basic problem of non-uniqueness

- Each datum is a “volumetric” response
- Data are

$$d_i = \sum_{j=1}^{M} G_{ij} m_j$$

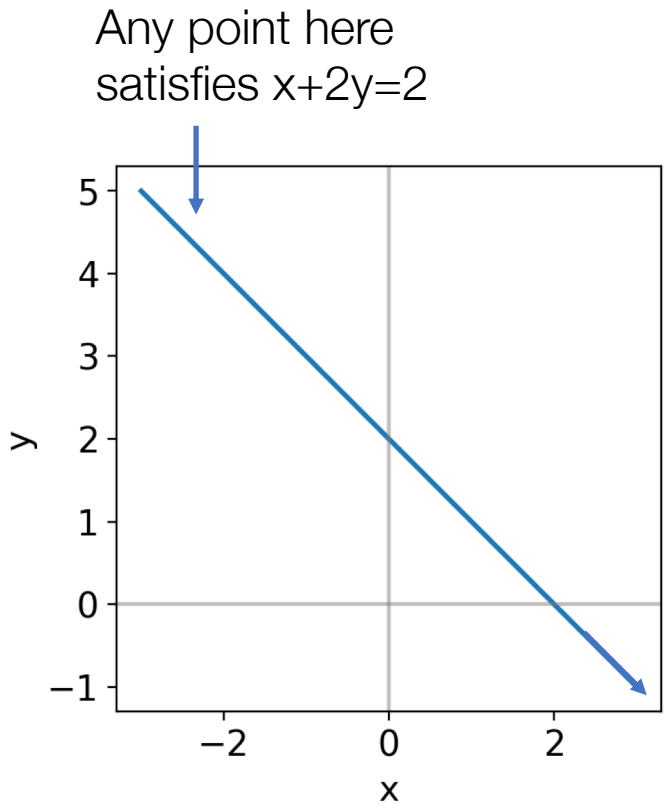
$$\mathbf{d} = \mathbf{G}\mathbf{m} \qquad \mathbf{G} : (N \times M)$$

- In the app
  - M=100
  - N=20
  - So, M>N (underdetermined problem) → infinitely many solutions

Questions to consider

- Consider the simple problem that involves two unknowns: m_1 and m_2
  - We have one datum: $m_1 + 2m_2 = 2$

- What is the value of m_1 and m_2?
  - (1, 0.5)
  - (2,0)
  - ….



Possible norms

- Smallness

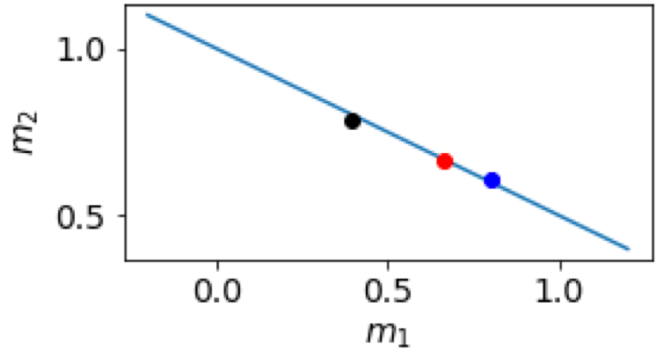
$$\phi_m = \|\mathbf{m}\|^2 = \sum m_j^2$$

- Smallness with reference

$$\phi_m = \|\mathbf{m} - \mathbf{m}_{\text{ref}}\|^2 = \sum_j (m_j - m_{\text{ref }j})^2$$

- Smoothness

$$\phi_m = \|\frac{d\mathbf{m}}{d\mathbf{x}}\|^2 = \sum_j (m_{j+1} - m_j)^2$$



$$\phi_m = m_1^2 + m_2^2$$

$$\rightarrow m = (0.392, 0.784)$$

$$\phi_m = (m_1 - m_{\text{ref}})^2 + (m_2 - m_{\text{ref}})^2$$

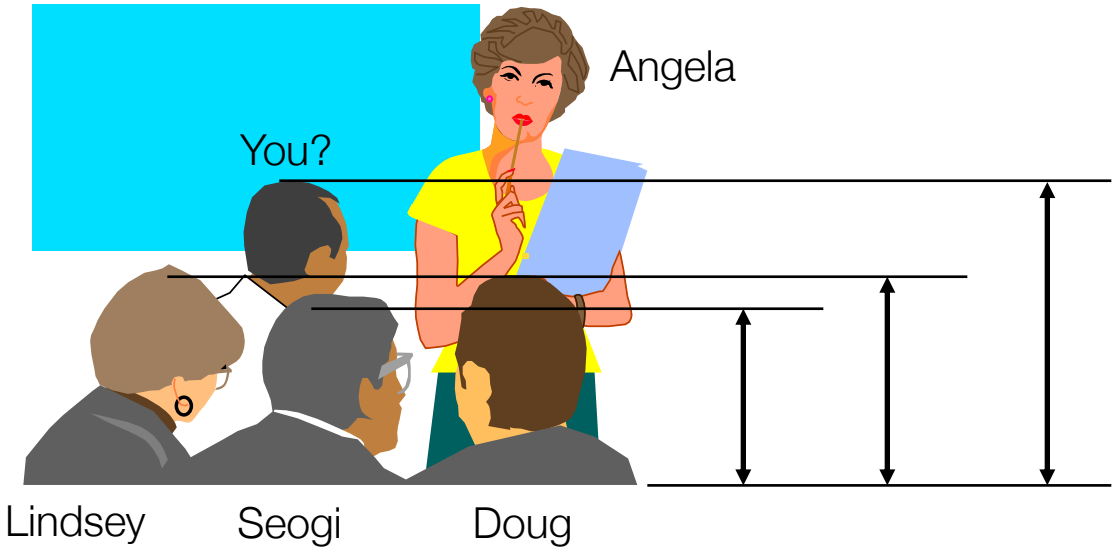
$$\rightarrow m = (0.8, 0.61) \qquad m_{ref} = 1$$

$$\phi_m = (m_1 - m_2)^2$$

$$\rightarrow m = (2/3, 2/3)$$

How to pick ”one” from ”many”

- Define: ruler to measure size (norm) of models
- Choose the “smallest”
- e.g. height



Model norms: Generally finding functions

Smallest model: $$\phi_m = \int m^2 dx$$

Smallest with reference: $$\phi_m = \int (m - m_{ref})^2 dx$$

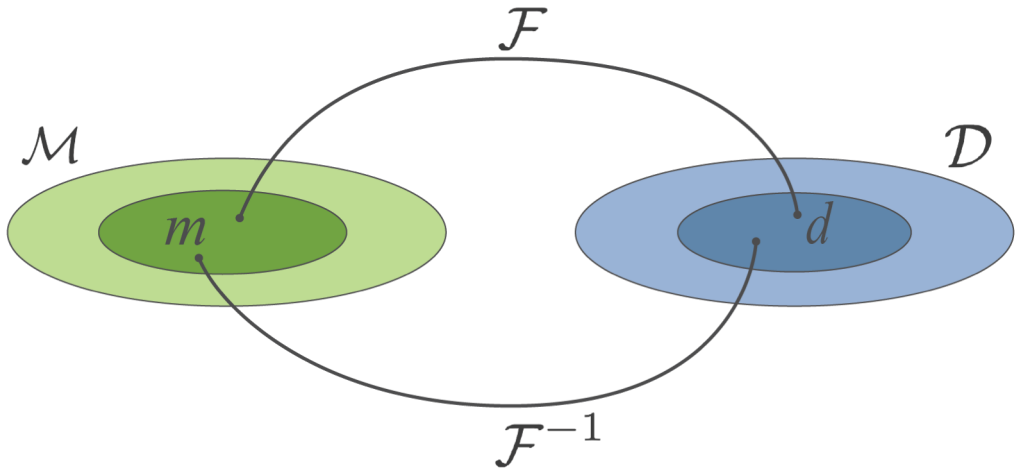
Smoothest model: $$\phi_m = \int \left(\frac{dm}{dx}\right)^2 dx$$

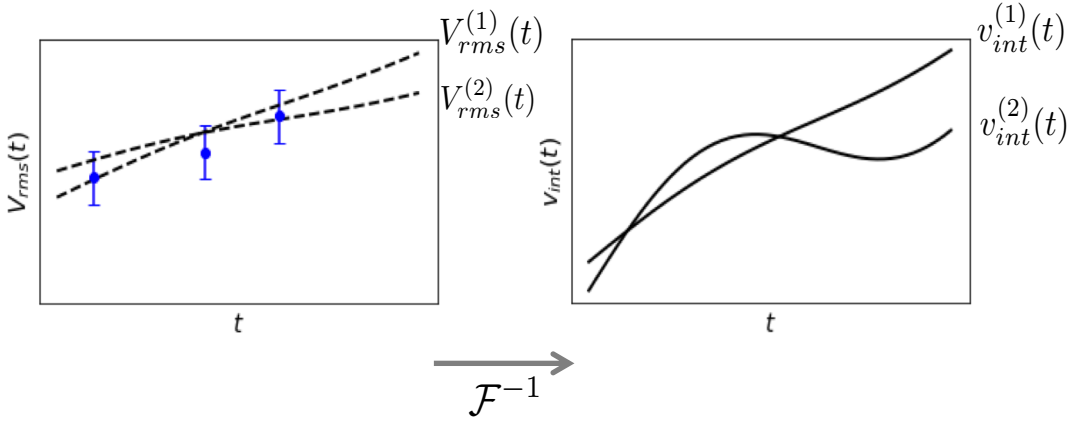
Combination: $$\phi_m = \alpha_s \int (m - m_{ref})^2 dx + \alpha_x \int \left(\frac{dm}{dx}\right)^2 dx$$

Discretize: $$\phi_m = \alpha_s \|\mathbf{W}_s(\mathbf{m} - \mathbf{m}_{ref})\|_2^2 + \alpha_x \|\mathbf{W}_x(\mathbf{m})\|_2^2$$

Summary

- Norms provide ways to select a specific solution
- Also need to address data errors and a misfit criterion





Inversion using Misfit criterion

Forward modelling	$\mathbf{Gm} = \mathbf{d}$
Data	$\mathbf{d}^{obs} = \mathbf{d} + \boldsymbol{\epsilon}$
Noise	$\boldsymbol{\epsilon}$

- Gaussian errors with standard deviation, ϵ_j
- Misfit measure: $\phi_d = \sum_{j=1}^{N} \left(\frac{d_j - d_j^{obs}}{\epsilon_j} \right)^2$
- Expected value of ϕ_d is $E[\phi_d] = N$
- Data are fit when $\phi_d \simeq \phi_d^*$

ϕ_d^*: target misfit

$\phi_d^* = N$

Combining misfit and model norm

- Goal for our inverse problem: find the model *m* that
  - produces an acceptable misfit ($\phi_d < \phi_d^*$)
  - minimizes the model norm, ϕ_m

- Re-cast as an optimization:

$$\text{minimize} \quad \phi_d + \beta\phi_m \qquad\qquad \text{where } 0 < \beta < \infty$$

- β : trade-off (Tikhonov) parameter

The role of β

Analogy: an optimization problem with two requirements

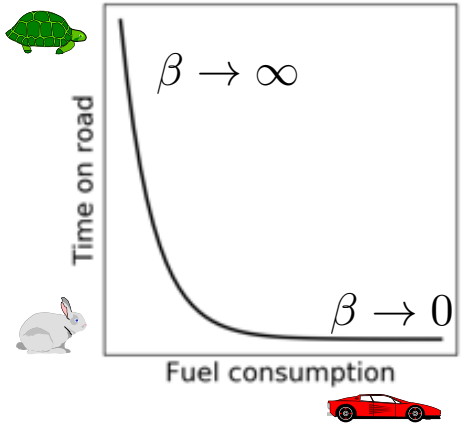
- Travelling from A to B
  - minimize **time** taken
  - minimize **fuel** consumption

$$\phi = \text{time} + \beta \cdot \text{fuel}$$

  - both time and fuel consumption are functions of speed

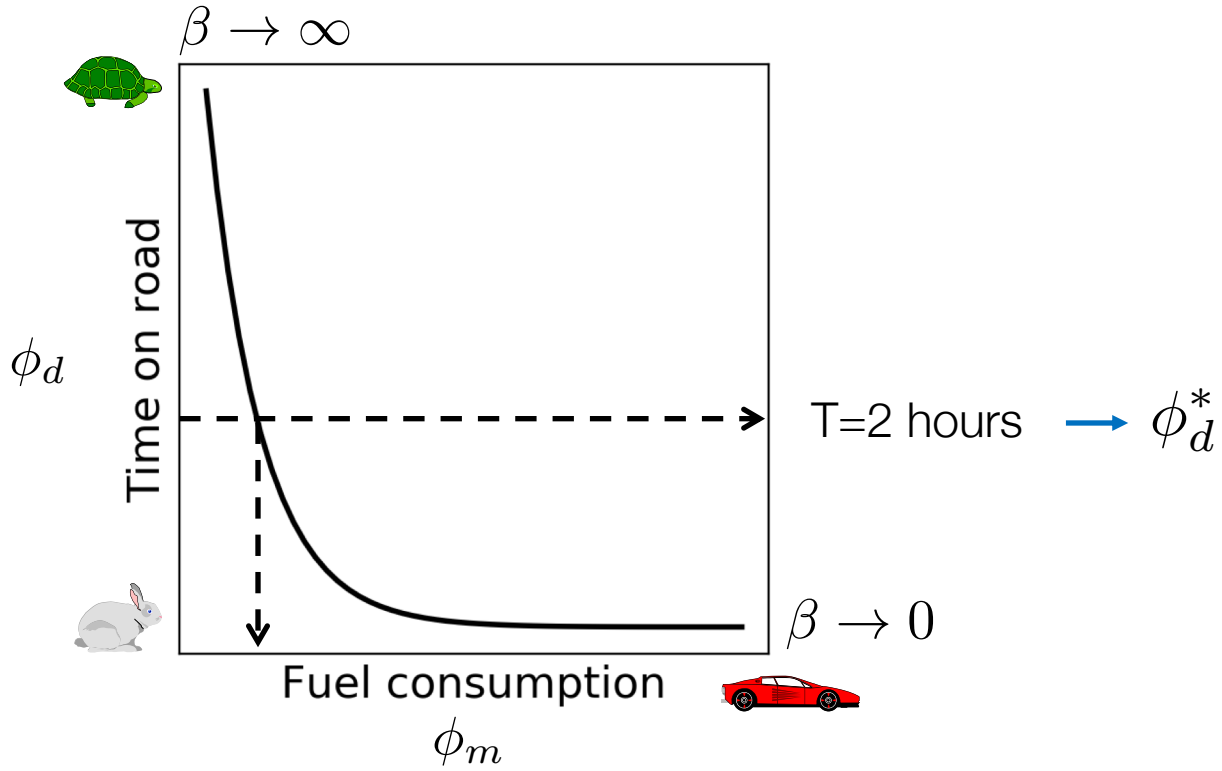
- $\beta = 0$: minimize time (regardless of fuel)

- $\text{large } \beta$: minimize fuel (but still get there)



The role of β

- A typical problem might be:
  - Minimize fuel consumption
  - Subject to getting there in 2 hours



$$\phi = \text{time} + \beta \cdot \text{fuel}$$

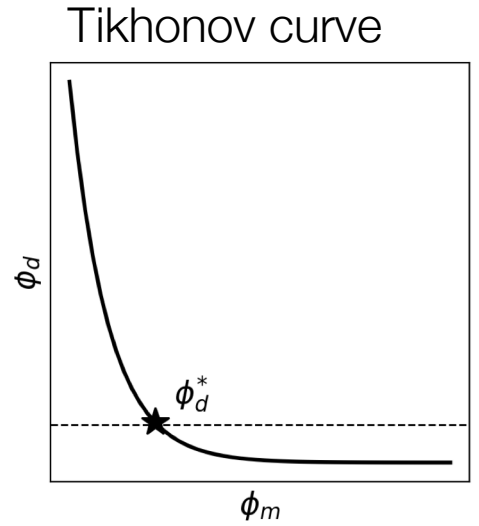
ϕ_d $\quad$ ϕ_m

The role of β : managing misfit

- Our inverse problem
  - Find the model (m): $\text{minimize} \quad \phi_d + \beta\phi_m$
  - Which beta to use?
  - If standard deviations of data are known,

$$E[\phi_d] = N$$

  - Desired misfit is $\phi_d^* \simeq N$
  - Choose β so that $\phi_d(m) = \phi_d^*$



Inversion app (demo)

mode ◉ Run ○ Explore

m0 0.00

mref 0.00

percentage 5

floor 0.02

chifact 1

data obs/pred | misfit

beta_min 0.0001

beta_max 100

} set range of β

n_beta 81

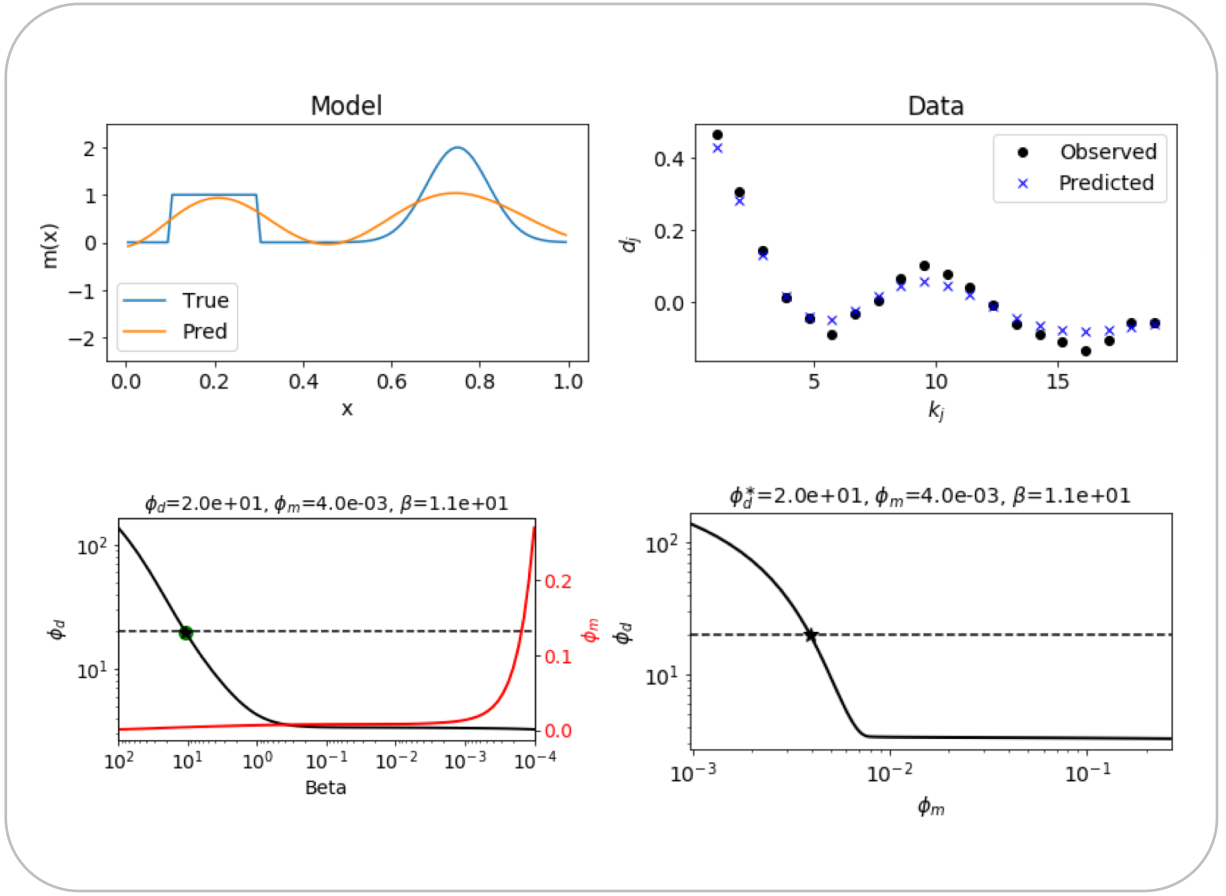
alpha_s 0.01

alpha_x 0

option misfit | tikhonov

i_beta 0

scale linear | log



β is the trade-off parameter

- Solve

$$\text{minimize} \quad \phi_d + \beta\phi_m$$

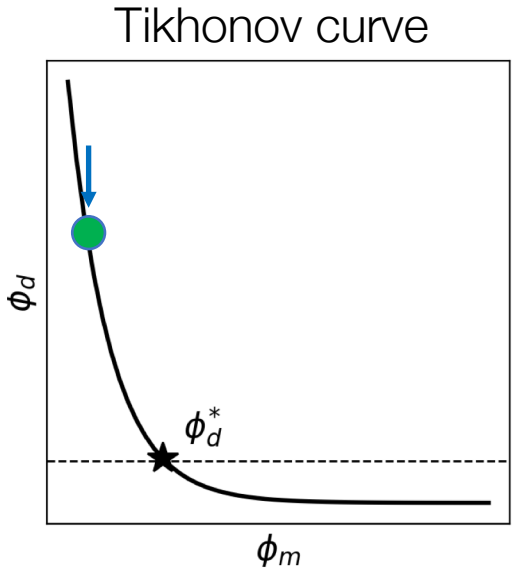
- β too large → **underfitting**
  - Structural information lost

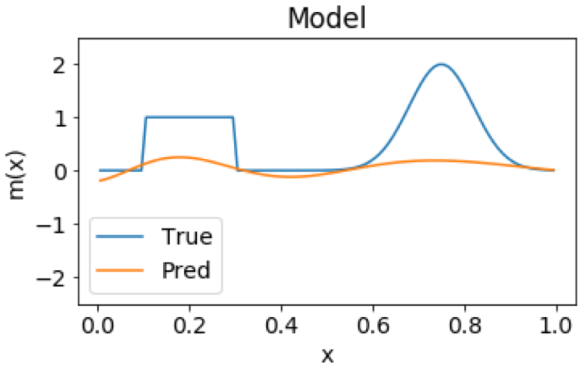
- β too small → overfitting the data
  - Noise becomes imaged as structure

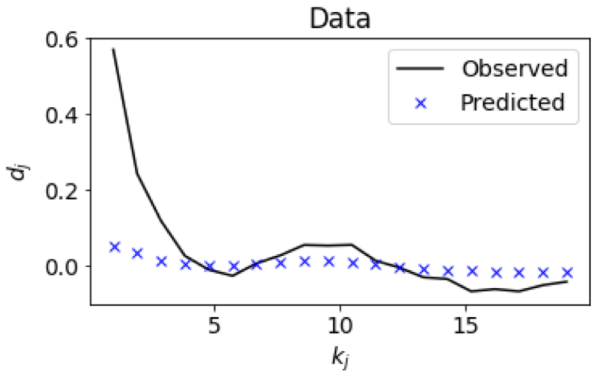
- β just right → optimal fit

$$\phi_d(m) \simeq N$$

  - Best estimate of a model which adequately re-creates the observations







β is the trade-off parameter

- Solve

$$\text{minimize} \quad \phi_d + \beta\phi_m$$

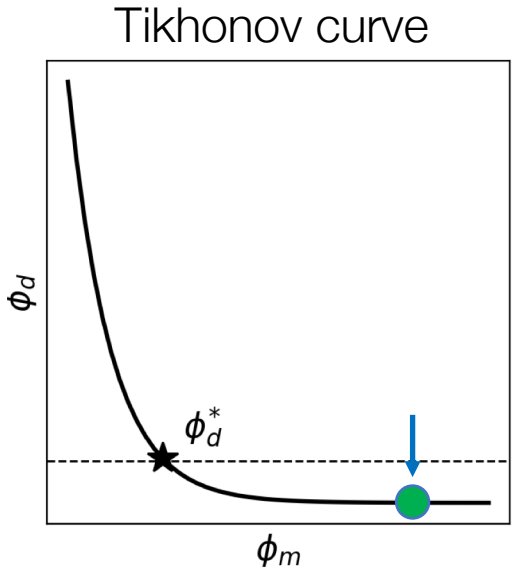
- β too large → underfitting
  - Structural information lost

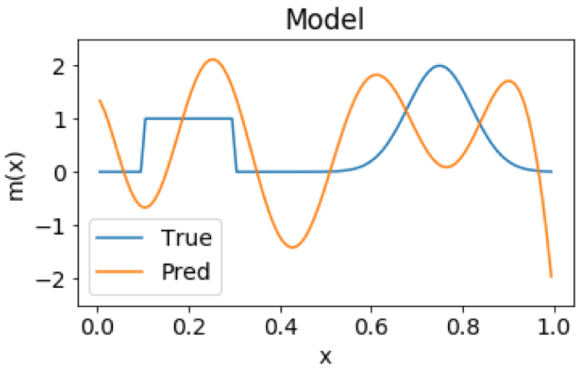
- β too small → **overfitting** the data
  - Noise becomes imaged as structure

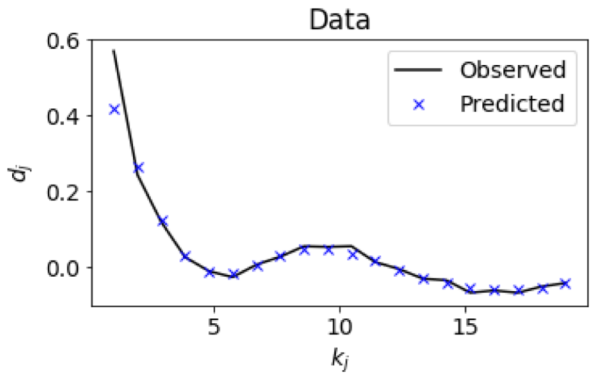
- β just right → optimal fit

$$\phi_d(m) \simeq N$$

  - Best estimate of a model which adequately re-creates the observations







β is the trade-off parameter

- Solve

$$\text{minimize} \quad \phi_d + \beta\phi_m$$

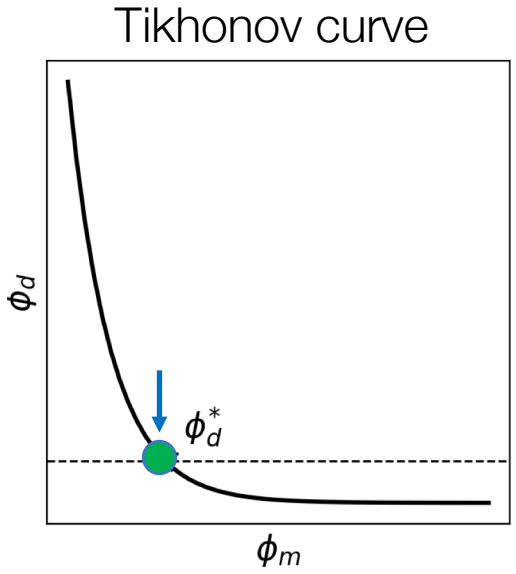
- β too large → underfitting
  - Structural information lost

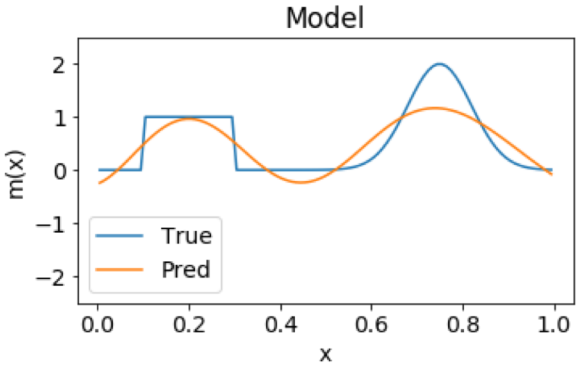
- β too small → overfitting the data
  - Noise becomes imaged as structure

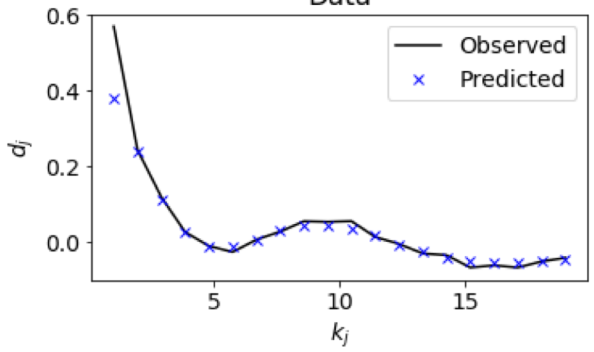
- β just right → **optimal fit**

$$\phi_d(m) \simeq N$$

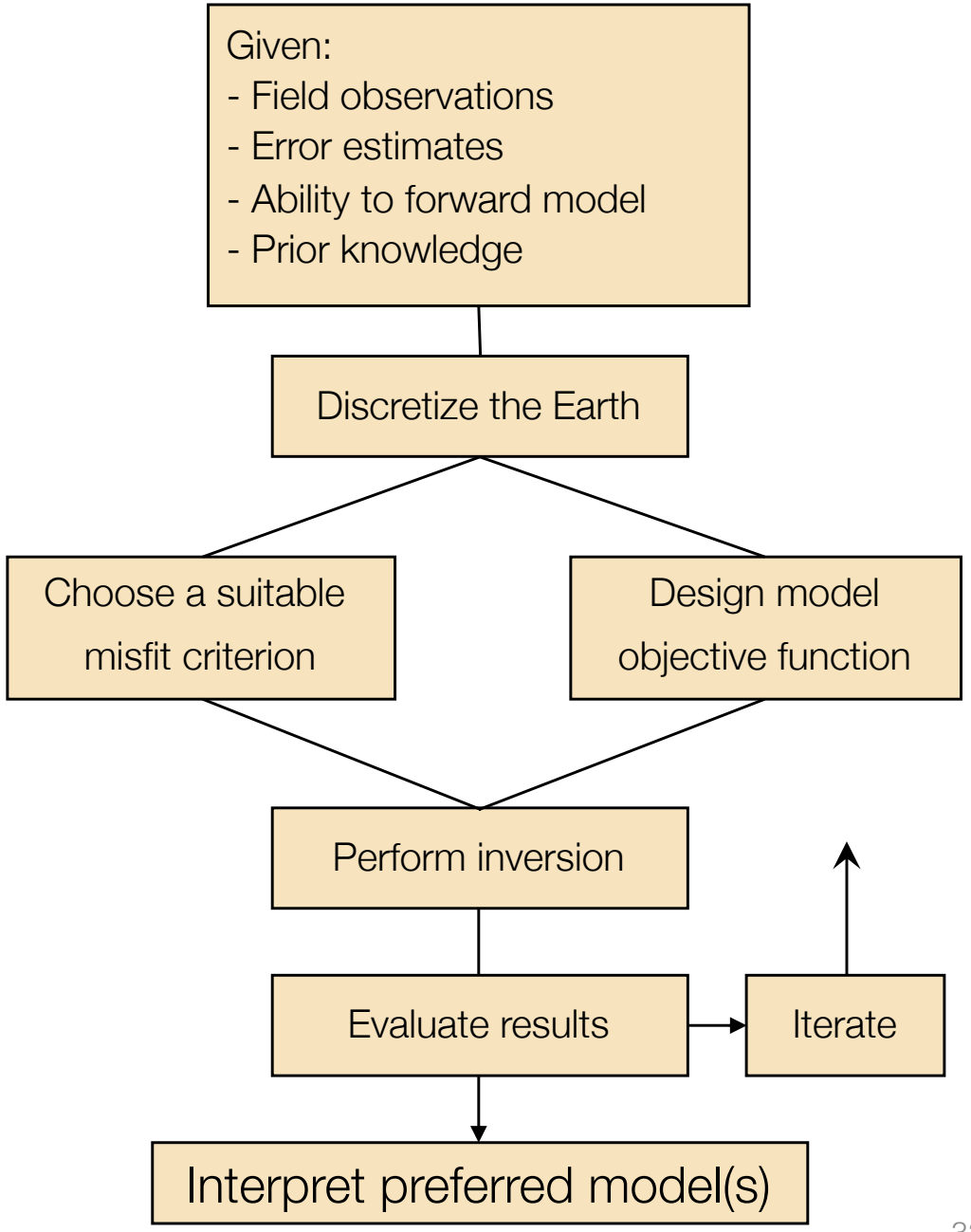
  - Best estimate of a model which adequately re-creates the observations





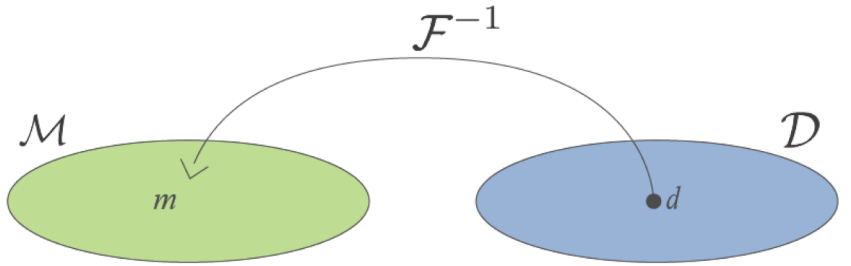


Flow chart for inverse problem



Summary

- Solving linear inverse problem



$$\text{minimize} \quad \phi_d + \beta\phi_m$$

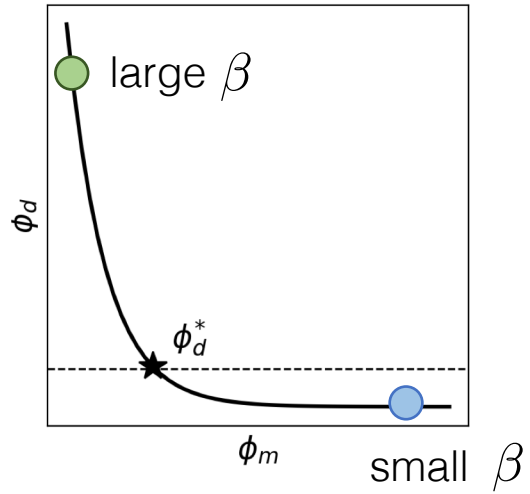
Data misfit

$$\phi_d = \sum_{j=1}^{N} \left(\frac{d_j - d_j^{obs}}{\epsilon_j} \right)^2 \qquad \begin{cases} \phi_d \simeq \phi_d^* \\ \phi_d^* = N \end{cases}$$

Model objective function
(or regularization)

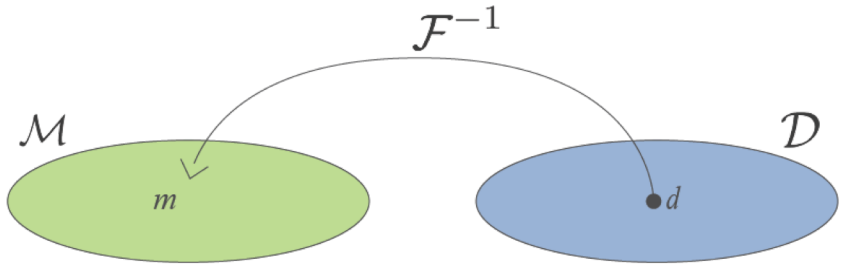
$$\phi_m = \alpha_s \int (m - m_{ref})^2 dx + \alpha_x \int \left(\frac{dm}{dx} \right)^2 dx$$

Trade-off parameter, β



Summary

- Solving linear inverse problem



$$\text{minimize} \quad \phi_d + \beta\phi_m$$

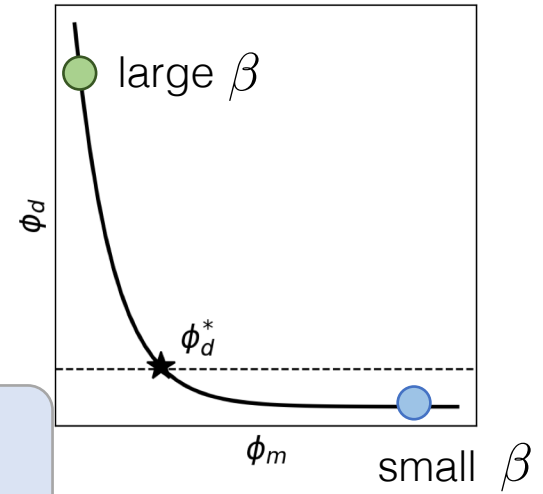
Trade-off parameter, β

Data misfit

$$\phi_d = \sum_{j=1}^{N} \left(\frac{d_j - d_j^{obs}}{\epsilon_j} \right)^2 \qquad \begin{cases} \phi_d \simeq \phi_d^* \\ \phi_d^* = N \end{cases}$$

Model objective function (or regularization)

$$\phi_m = \alpha_s \int (m - m_{ref})^2 dx + \alpha_x \int \left(\frac{dm}{dx} \right)^2 dx$$



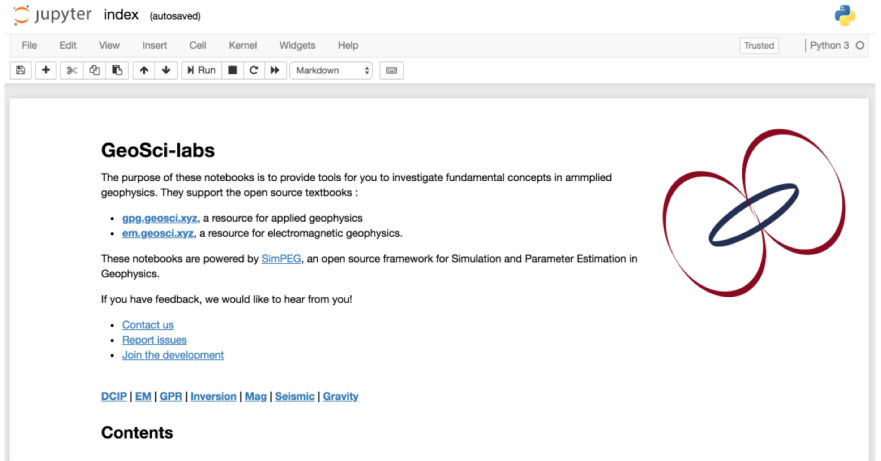
How do we solve this optimization problem?

Outline

- Statement of the inverse problem
- App for a linear problem
- Data misfit
- Model norms and their importance.
- Tikhonov approach to inversion
- A flow chart to keep you on track

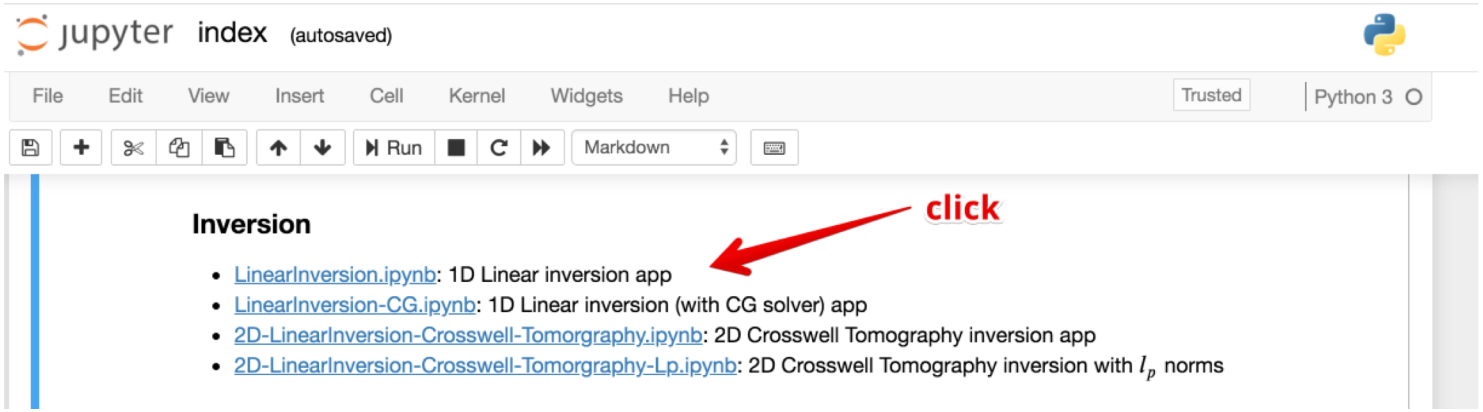
Homework problem

- Accessing app
  - Link: https://mybinder.org/v2/gh/geoscixyz/geosci-labs/master?filepath=notebooks%2Findex.ipynb



  - Wait until you see index page

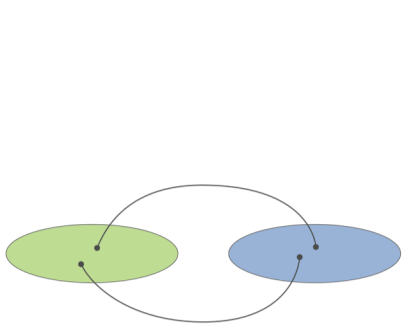
  - Then



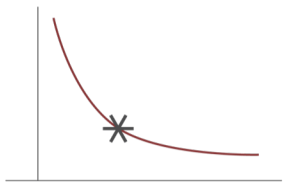
Homework problem

- Question:
  - Use default parameters in LinearInversion.ipynb
  - Except data error (%) =0 (keep floor = 0.01)
  - What is the value of β that corresponds to the inflection point of the Tikhonov curve (corner of the L-curve)

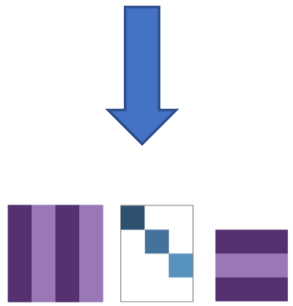
Next up …



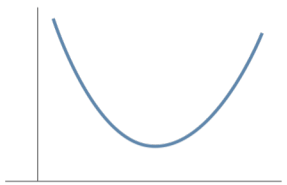
Overview



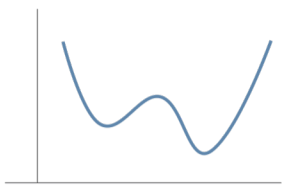
Tikhonov



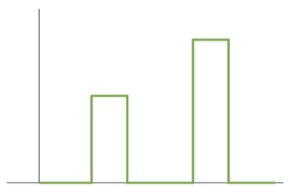
SVD



Linear



Nonlinear



Lp-norms

Field scale

The future