Probabilistic Inverse Theory

Lecture 4

W. Dębski

29.03.20223

Linear problem

$$\mathbf{m} \sim \mathbf{m}^{apr}; \qquad \bar{\mathbf{d}} = \mathbf{G} \cdot \mathbf{m}^{apr}$$

$$\mathbf{d} = \mathbf{G} \cdot \mathbf{m} \quad / - \bar{\mathbf{d}}$$

$$\mathbf{d} - \bar{\mathbf{d}} = \mathbf{G} \cdot (\mathbf{m} - \mathbf{m}^{apr})$$

$$\mathbf{G}^T \cdot (\mathbf{d} - \bar{\mathbf{d}}) = \mathbf{G}^T\mathbf{G} \cdot (\mathbf{m} - \mathbf{m}^{apr})$$

$$\mathbf{G}^T\mathbf{G} \implies \mathbf{G}^T\mathbf{G} + \lambda\mathbf{I}$$

$$\mathbf{m}^{est} = \mathbf{m}^{apr} + (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T \cdot \left(\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr}\right)$$

Linear problem - resolution analysis

$$\mathbf{m} = (m_1, m_2, \ldots \mathbf{m}_M)$$

- how accurately one can estimate m_i ?
- are m_i, m_j correlated ?

$$\mathbf{m}^{est} = \mathbf{m}^{apr} + (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T \cdot (\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr})$$

$$\mathbf{d}^{obs} = \mathbf{d}^{true}; \qquad \mathbf{d}^{true} = \mathbf{G} \cdot \mathbf{m}^{true}$$

$$\mathbf{m}^{est} - \mathbf{m}^{apr} = \mathbf{R}(\lambda) \cdot (\mathbf{m}^{true} - \mathbf{m}^{apr})$$

Linear problem - well defined problem

$$\det(\mathbf{G}^T \cdot \mathbf{G}) \neq 0 \longrightarrow \lambda = 0$$

$$\mathbf{R} = \frac{\mathbf{G}^T\mathbf{G}}{\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I}}$$

$$\mathbf{R} = \mathbf{I}$$

$$\mathbf{m}^{est} = \mathbf{m}^{true}$$

Linear problem - under-determined problem

$$\det(\mathbf{G}^T \cdot \mathbf{G}) = 0 \longrightarrow \lambda \neq 0$$

$$\mathbf{R} = \frac{\mathbf{G}^T\mathbf{G}}{\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I}}$$

$$\mathbf{R} \neq \mathbf{I}$$

$$\mathbf{m}^{est} \neq \mathbf{m}^{true}$$

Linear problem - inversion errors

$$\mathbf{m}^{apr} = \mathbf{m}^{true}; \qquad \mathbf{d}^{obs} = \mathbf{d}^{true} + \delta$$

$$\mathbf{m}^{est}(\lambda) = \mathbf{m}^{true} + (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T \cdot \left(\mathbf{d}^{obs} - \mathbf{d}^{true}\right)$$

$$\mathbf{m}^{est} = \mathbf{m}^{true} + (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T \cdot \delta$$

$$\boldsymbol{\Delta}\mathbf{m}^{\mathbf{est}} = (\mathbf{G}^T\mathbf{G} + \lambda\mathbf{I})^{-1}\mathbf{G}^T \cdot \delta$$

Linear problem - dependences on m^{apr} (1D)

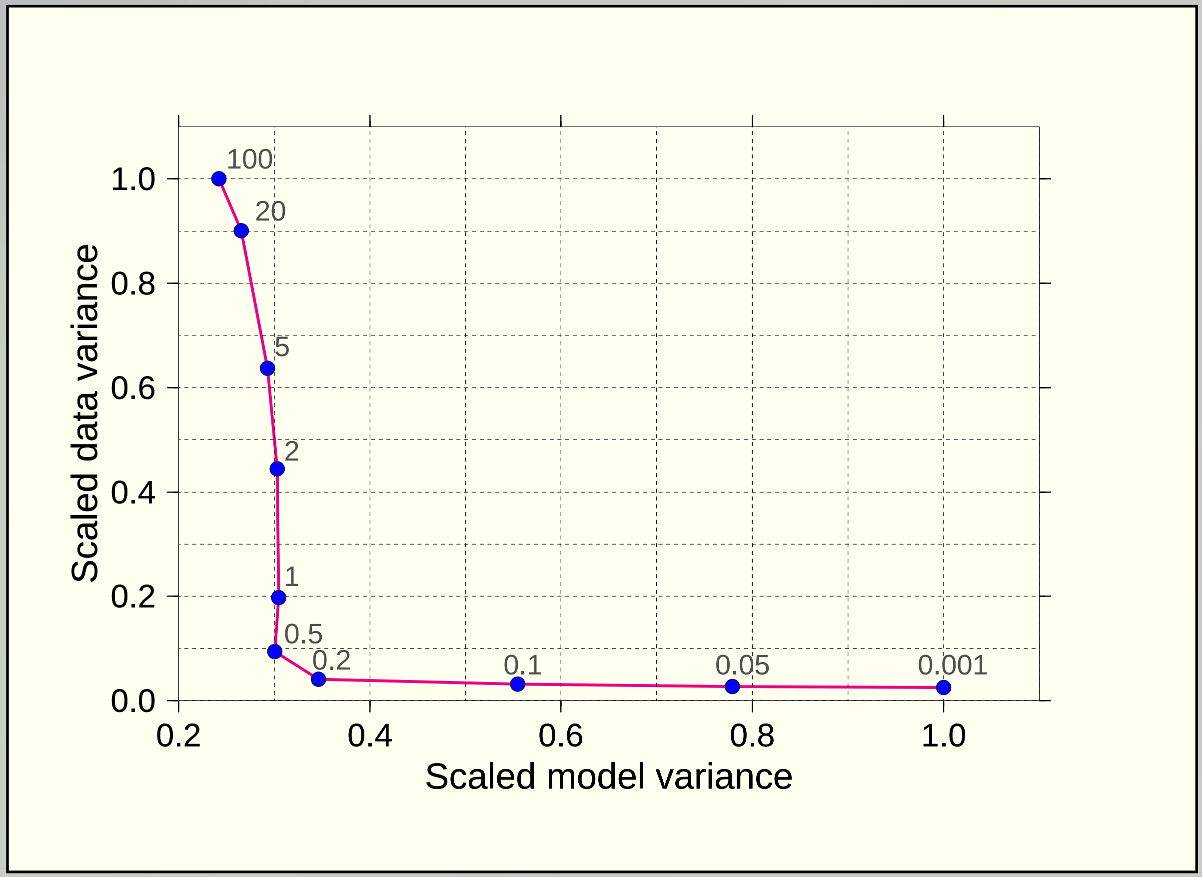
$$d = g \cdot m \qquad d^{obs} = d^{true} + \delta \qquad \lambda = \epsilon g^2$$

$$\Delta m^{a/e} = m^{apr/est} - m^{true} \quad \Delta d^{a/e} = d^{apr/est} - d^{true}$$

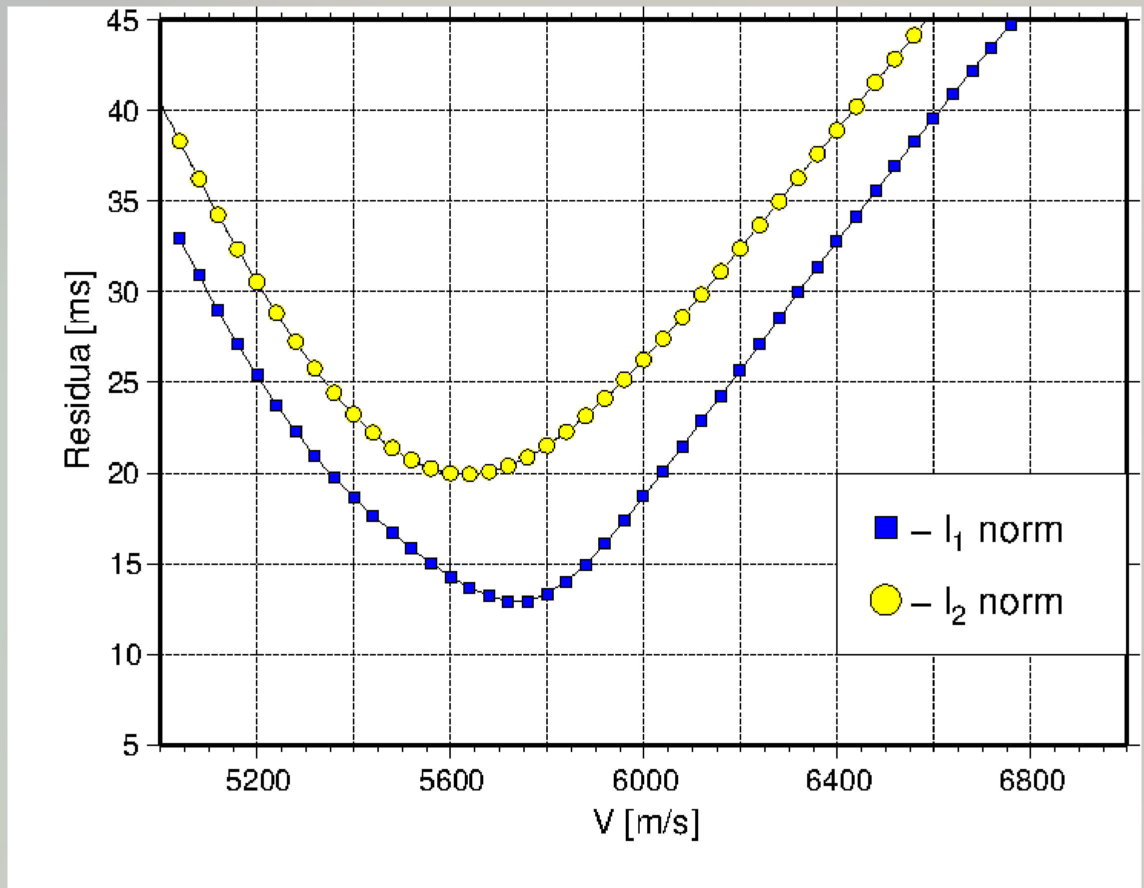
$$\Delta m^e = \frac{\epsilon}{1+\epsilon}(\Delta m^a) + \frac{1}{1+\epsilon}\left(\frac{\delta}{g}\right)$$

$$\Delta d^e = \frac{\epsilon}{1+\epsilon}(\Delta d^a) + \frac{\delta}{1+\epsilon}$$

Linear problem - selecting λ



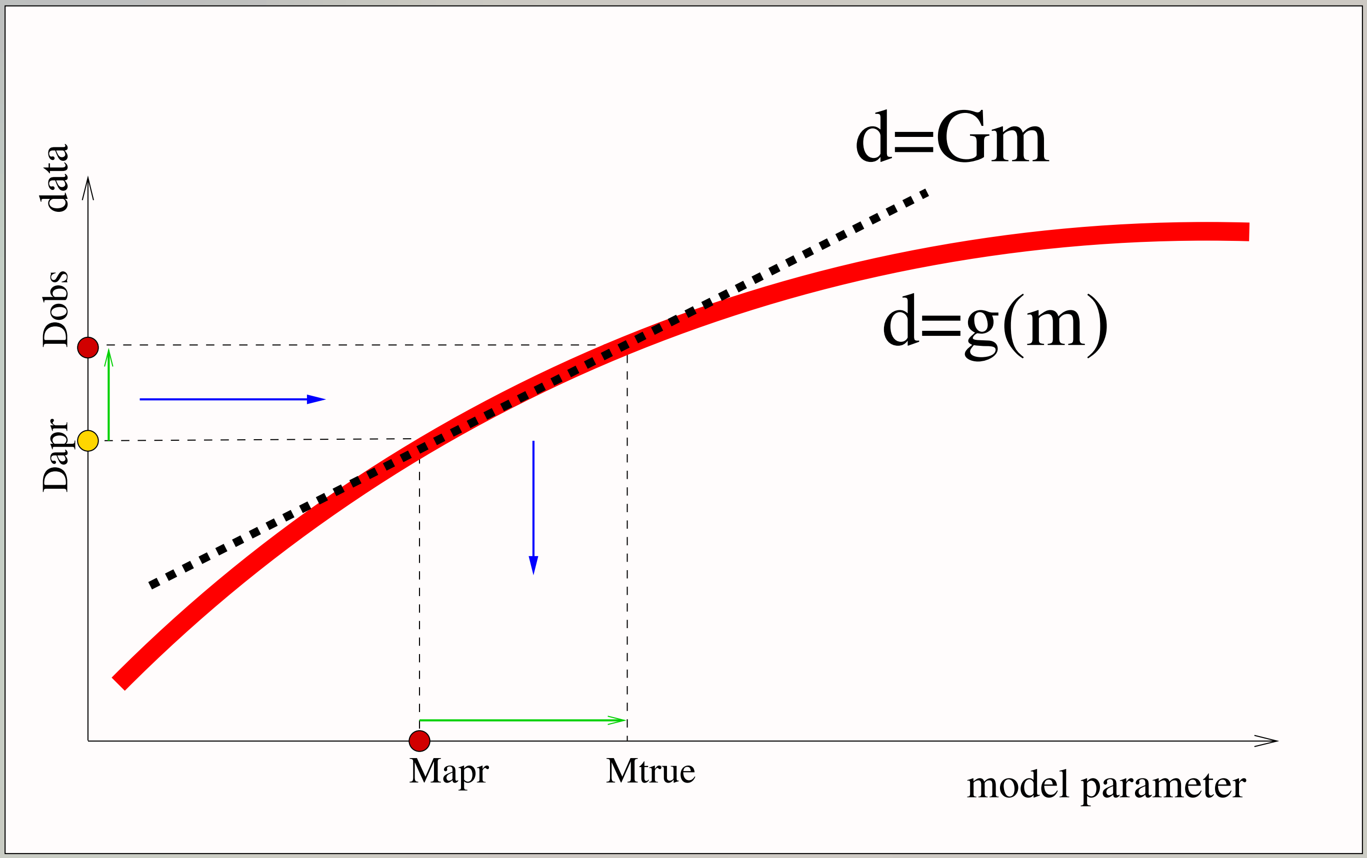
Linear problem - selecting $\mathbf{m}^{apr}$



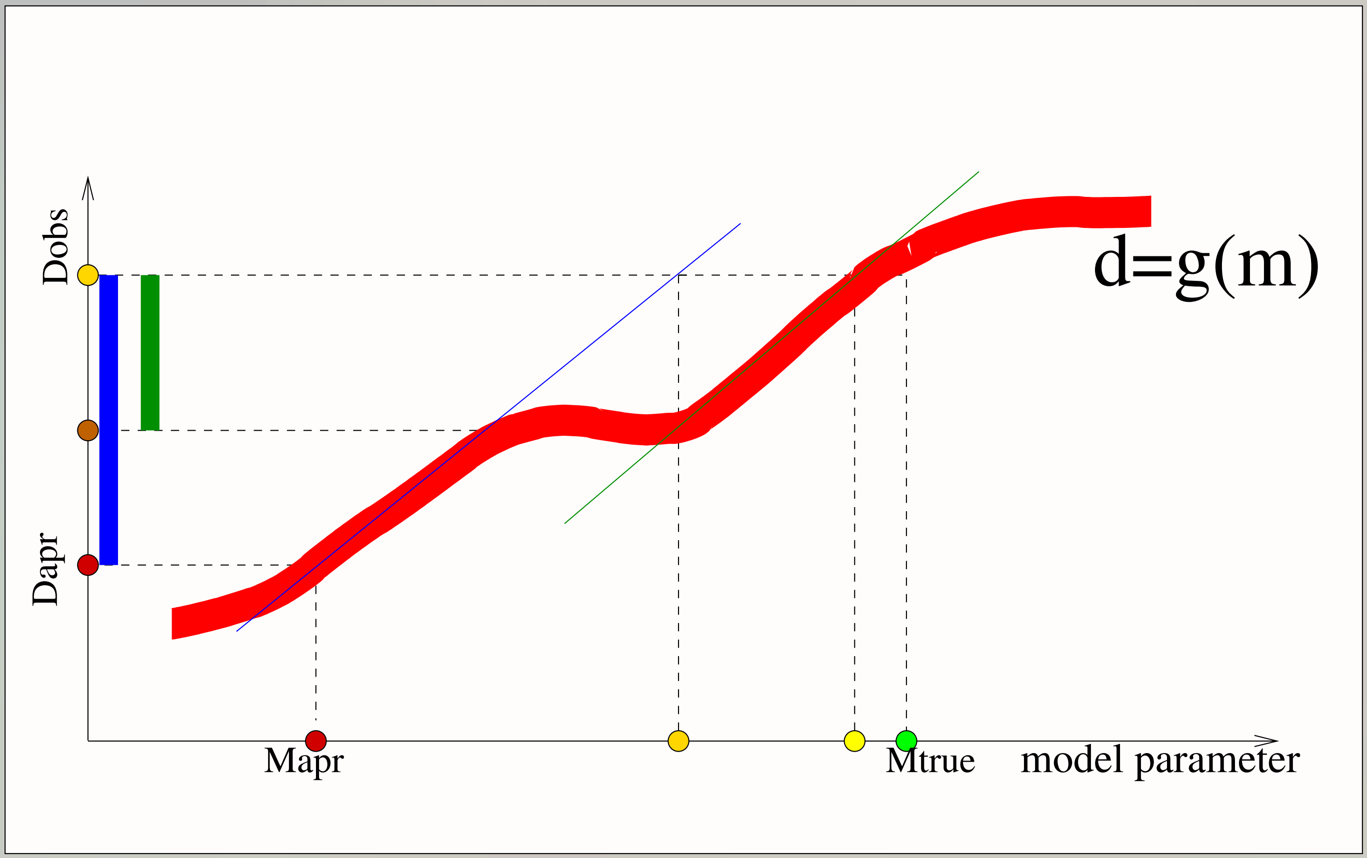
Algebraic approach - features

- ✦ simplicity
- ✦ very fast (huge inverse problems)
- ✦ only linear problem
- ✦ lack of robustness

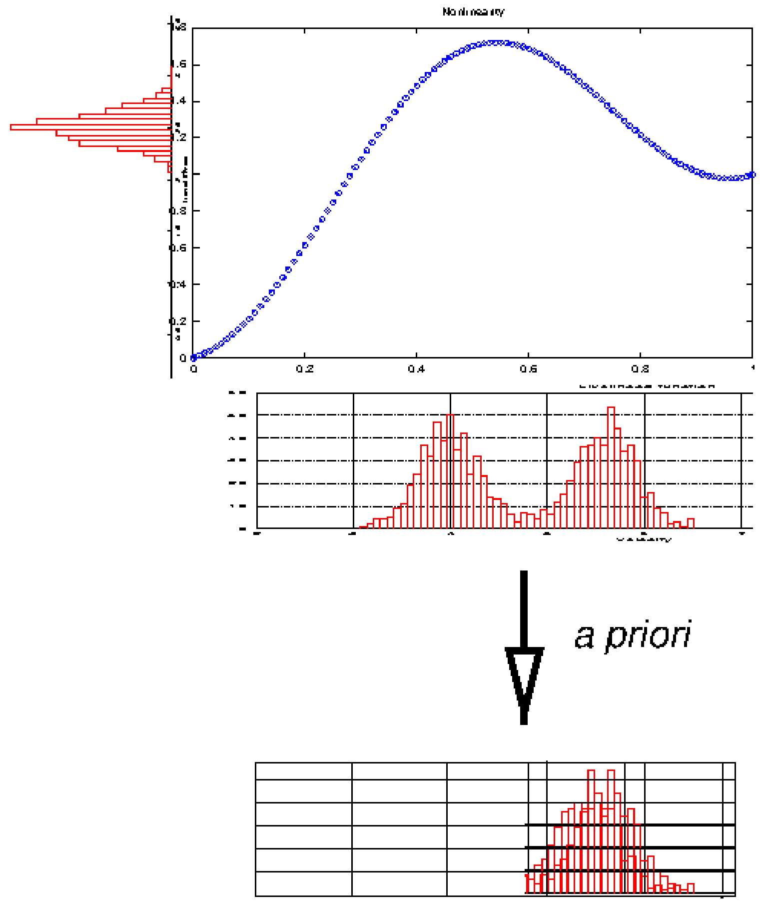
Algebraic approach - back projection



Nonlinear problems



Nonlinear problems



Optimization approach - generalization

Linearity of the problem

$$\mathbf{d} = \mathbf{G} \cdot \mathbf{m}$$

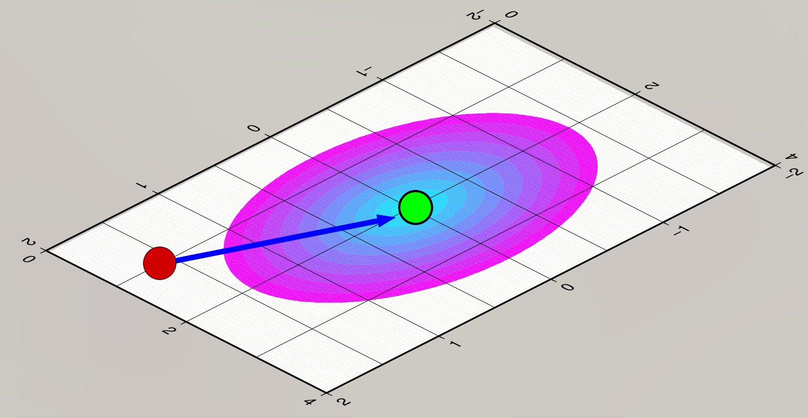
has allowed to move from *a priori* to *true* model

$$\mathbf{m}^{apr} \Longrightarrow \mathbf{m}^{true} :$$

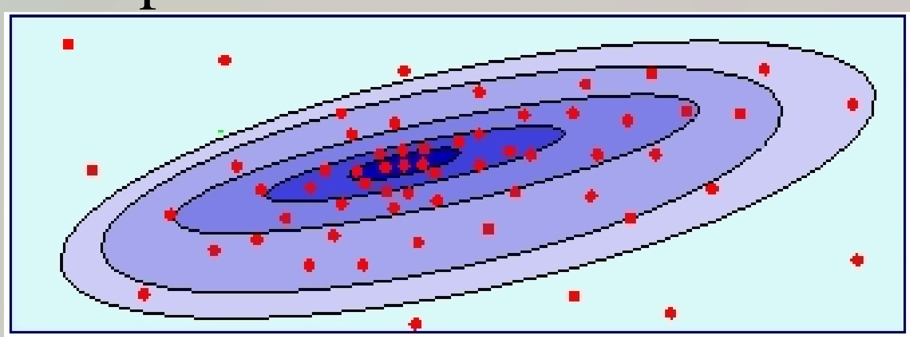
in a “single step”

Optimization approach - generalization

$$\begin{array}{ccc} \mathcal{D} & \Longrightarrow & \mathcal{D} \\ \uparrow & & \downarrow \\ \mathcal{M} & & \mathcal{M} \end{array}$$



Search over the $\mathcal{M}$ space for a model which best reproduces $\mathbf{d}^{obs}$



Optimization approach - direct search for best m

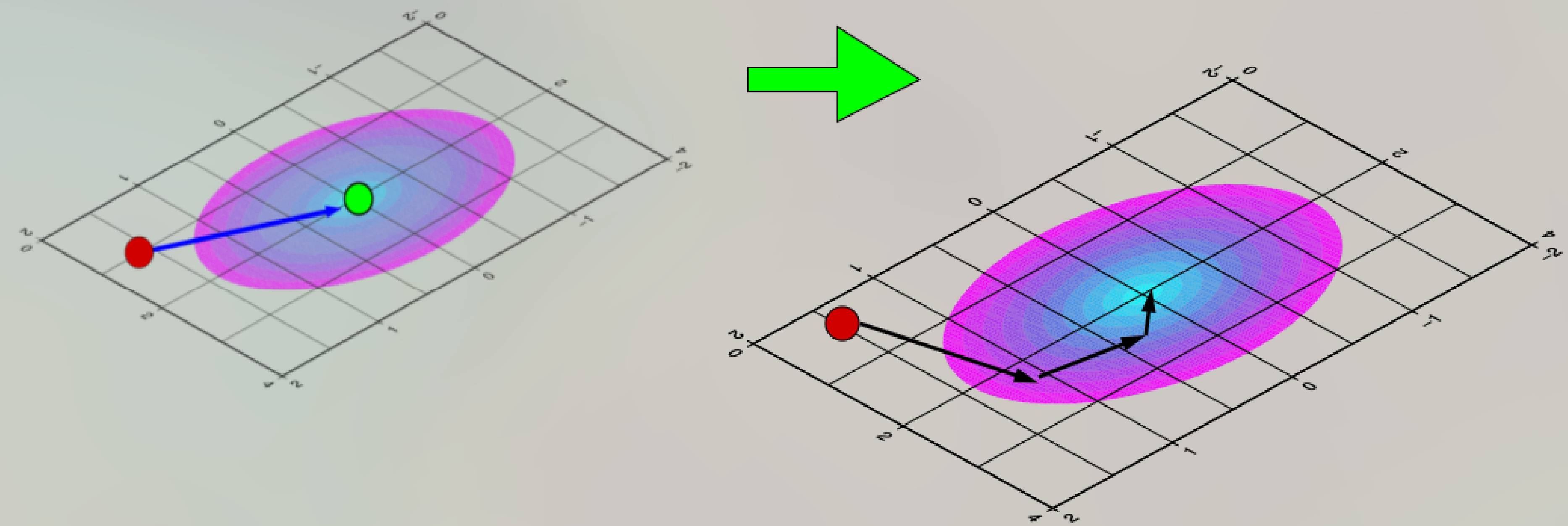
Generalization

$$S(\mathbf{m}) = ||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m})||$$

solution: search for $\mathbf{m}^{ml}$ minimizing $S(\mathbf{m})$

$$S(\mathbf{m}^{ml}) = \min$$

Inverse problem - optimization approach



Inverse problem $\equiv$ minimization cost function

Inverse problem - optimization approach

$$||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{est})|| = \min$$

more generally

$$||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{ml})||_{\mathcal{D}} + ||\mathbf{m}^{ml} - \mathbf{m}^{apr}||_{\mathcal{M}} = \min$$

Optimization approach - features

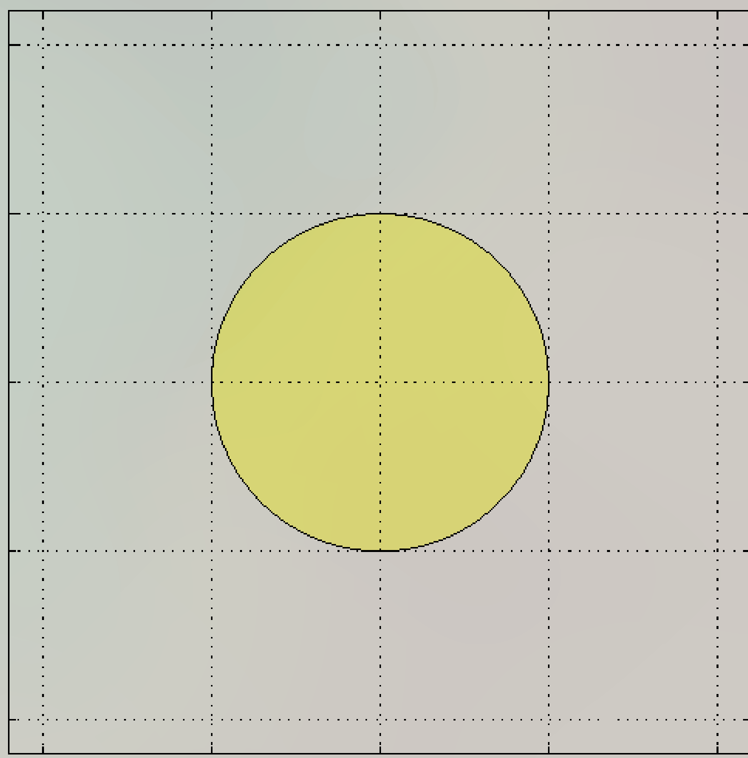
- fully nonlinear method
- physical meaning of regularization
- variety of existing optimization method
- problem with error estimation
- How to define misfit function $S(\mathbf{m})$?
- Regularization not needed but ...
- is solution unique ?

Norms

l_1	$\|\|\mathbf{d}\|\| = \sum_i \left\| \frac{d^i}{C^i} \right\|$	Absolute value norm
l_2	$\|\|\mathbf{d}\|\| = \sum_{ij} d^i C^{ij} d^j$	Gaussian norm
l_c	$\|\|\mathbf{d}\|\| = \sum_i \log\left(1 + \left(\frac{d^i}{C^i}\right)^2\right)$	Cauchy norm
l_m	$\|\|\mathbf{d}\|\| = \log\left(1 + \sum_i \left(\frac{d^i}{C^i}\right)^2\right)$	Modified Cauchy norm
l_s	$\|\|\mathbf{d}\|\| = \sum_i \log\left[\cosh\left(\frac{x_i}{C_i}\right)\right]$	Hyperbolic secant norm
l_p	$\|\|\mathbf{d}\|\| = \sqrt[p]{\sum_i \left(\frac{\|d^i\|}{C_i}\right)^p}$	generalized Gaussian norm

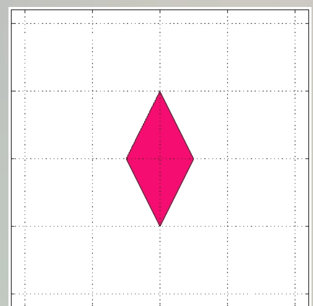
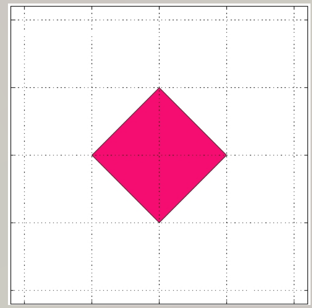
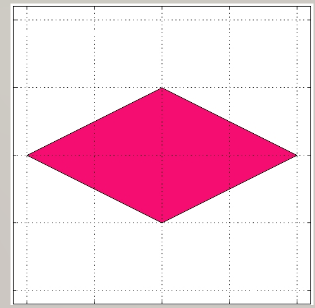
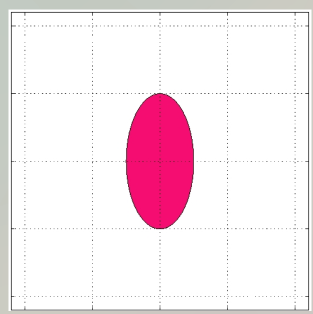
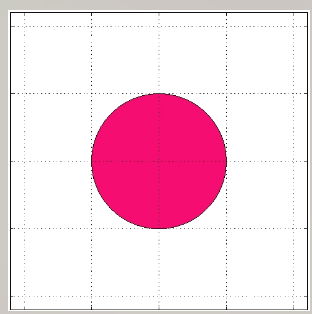
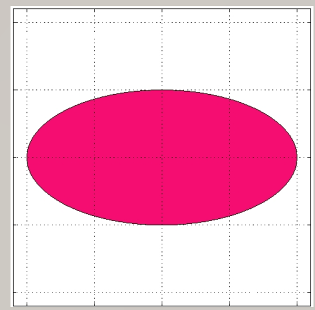
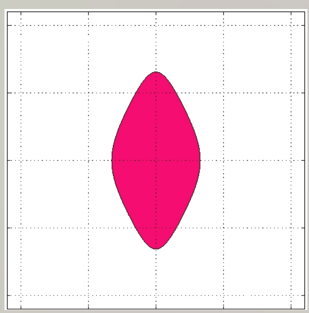
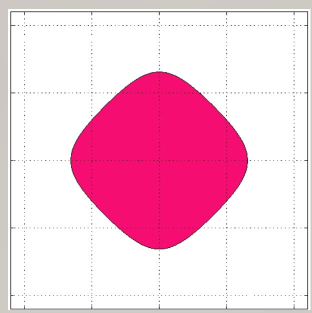
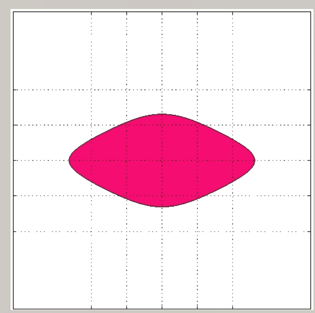
Different norms - circle

Circle of radius r and center in O by definition is formed by a set of points equally distanced from O



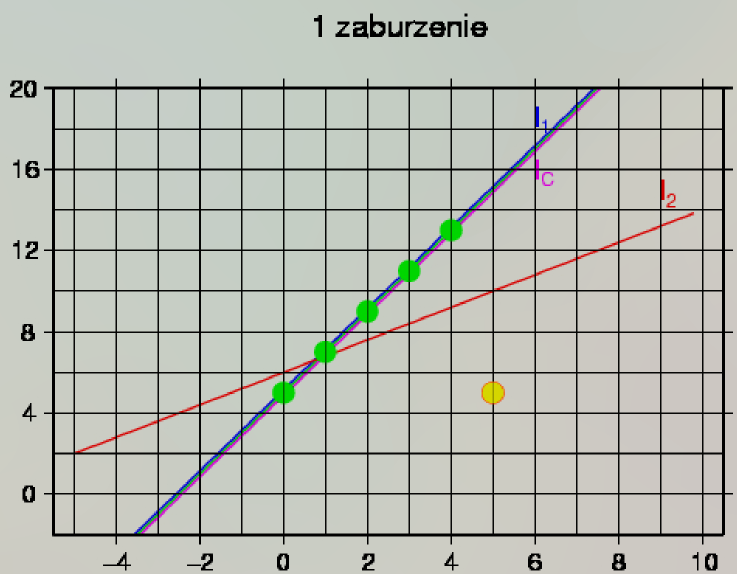
$$||\vec{\mathbf{r}}|| = 1$$

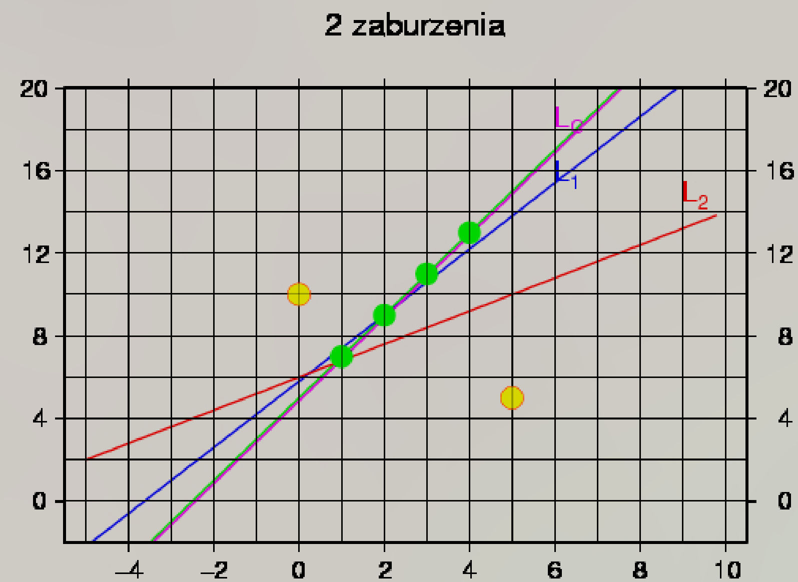
Different norms - circle

Norm	$c_1 = 2c_2$	$c_1 = c_2$	$c_1 = \frac{1}{2}c_2$
l_1			
l_2			
l_c			

Different norms - different solutions

$$\{(x_i, y_i)\}; \quad y = m_1 x + m_2 \Longrightarrow ||y_i - (m_1 x_i + m_2)|| = \min$$





Optimization approach - main steps

- selection *a priori* model $\mathbf{m}^{apr}$
- selection norms $||\cdot||_{\mathcal{D}}$ and $||\cdot||_{\mathcal{M}}$
- selection optimization algorithm
- run optimization
- post-optimization analysis (residua, resolution, etc.)

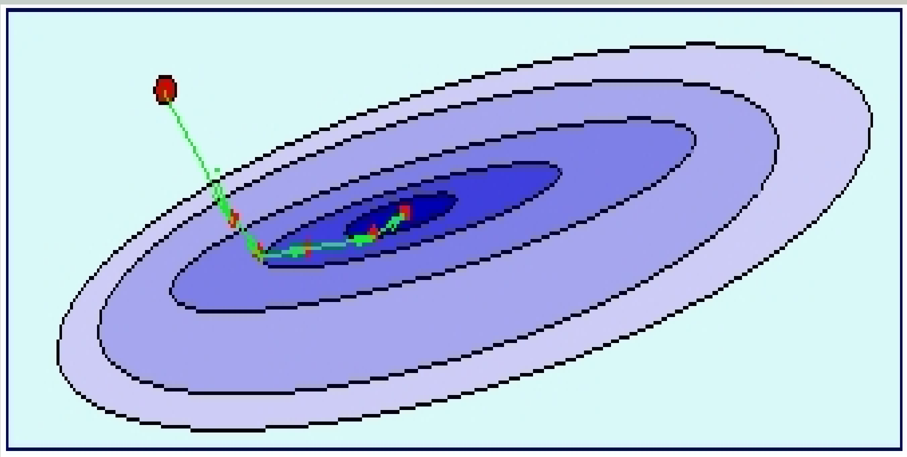
Optimization approach - norm selection

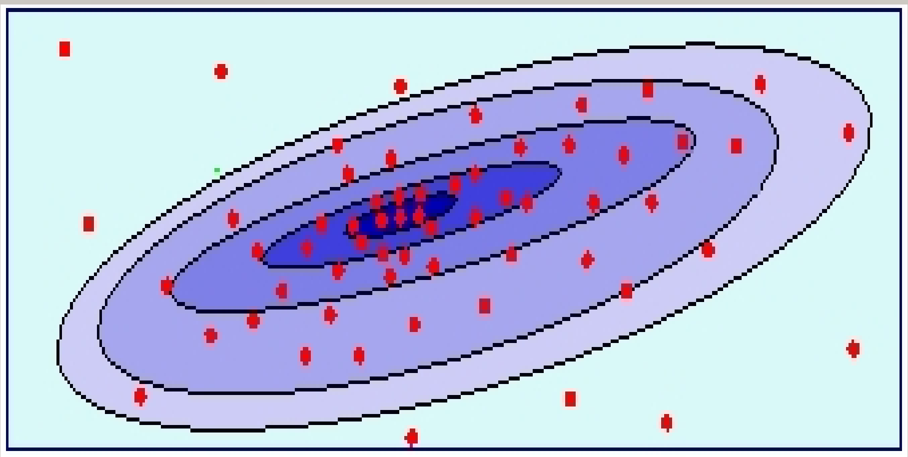
Data/model norm should follow expected features of measurement/modelling uncertainties. If Gaussian errors are expected the l_2 norm is the best choice as it is leading to the smallest errors (we shall discus it later on). If outliers in data are expected, the more robust norms like l_1 or l_C should be considered.

Optimization approach - optimizer selection

heuristic

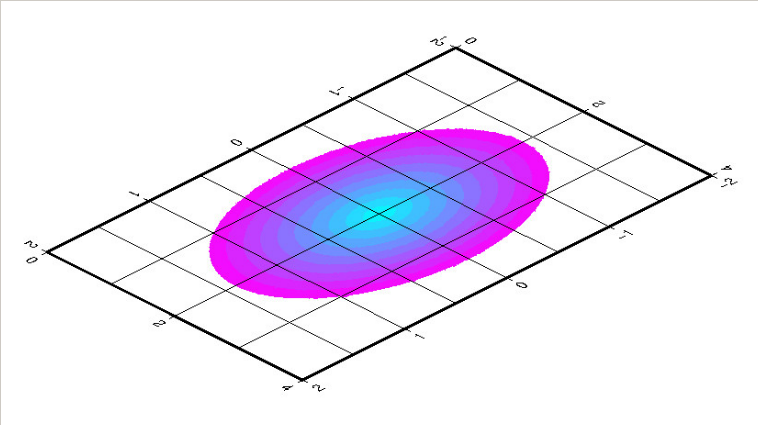
stochastic





Optimization approach - grid search method

```
Rmin = Rmax
for  m_1 = a_{min} :  a_{max}
 for m_2 = b_{min} :  b_{max}
  R = || d^{obs}\; - \;\d^{th}(\m) ||
  R = R +  || \m  - \ma ||
  if(R<Rmin)
    Rmin = R
    m_1^{est}\; = \;m_1
    m_2^{est}\; = \;m_2
  endif
 end
end
```



Optimization - *preconditioned steepest descent*

1. $\mathbf{m}_0$ - arbitrary,
2. $\mathbf{G}_n : G(\mathbf{m}) \approx \mathbf{G}_n \cdot (\mathbf{m} - \mathbf{m}_n)$
3. $\hat{\mathbf{S}}_0 \approx (\mathbf{I} + \mathbf{C}_M \mathbf{G}_n^T \mathbf{C}_D^{-1} \mathbf{G}_n)^{-1}$
4. $\gamma_n = \mathbf{C}_m \mathbf{G}_n^T \mathbf{C}_d^{-1} (G_n \mathbf{m}_n - \mathbf{d}^{obs}) + (\mathbf{m}_n - \mathbf{m}^{apr})$
5. $\phi_n = \hat{\mathbf{S}}_0 \gamma_n$
6. $\mathbf{b}_n = \mathbf{G}_n \phi_n$
7. $\mu_n = \frac{\gamma_n^t \mathbf{C}_m^{-1} \phi_n}{\phi_n^t \mathbf{C}_m^{-1} \phi_n + \mathbf{b}_n^t \mathbf{C}_d \mathbf{b}_n}$
8. $\mathbf{m}_{n+1} = m_n - \mu_n \phi_n$

Optimization approach - linear problem and l_2 norms

$$||\mathbf{x}|| = \mathbf{x}^T \mathbf{C}_d^{-1} \mathbf{x}$$

Misfit function

$$S(\mathbf{m}) = ||\mathbf{d}^{obs} - \mathbf{d}^{th}(\mathbf{m}^{ml})||_{\mathcal{D}} + ||\mathbf{m}^{ml} - \mathbf{m}^{apr}||_{\mathcal{M}}$$

$$\begin{aligned} S(\mathbf{m}) \;=\; & (\mathbf{d}^{obs} - \mathbf{G}\cdot\mathbf{m})^T\, \mathbf{C}_d^{-1}\, (\mathbf{d}^{obs} - \mathbf{G}\cdot\mathbf{m}) + \\ & (\mathbf{m} - \mathbf{m}^{apr})^T\, \mathbf{C}_m^{-1}\, (\mathbf{m} - \mathbf{m}^{apr}) \end{aligned}$$

Minimum condition:

$$\frac{dS(\mathbf{m})}{dm_i} = 0$$

Optimization approach - linear problem and l_2 norms

Solution:

$$\mathbf{m}^{est} \;=\; \mathbf{m}^{apr} + \mathbf{C}_p^{-1}\mathbf{G}^T\mathbf{C}_d^{-1}\cdot\left(\mathbf{d}^{obs} - \mathbf{G}\cdot\mathbf{m}^{apr}\right)$$

where

$$C_p = \left(\mathbf{G}^T\mathbf{C}_d^{-1}\mathbf{G} + \mathbf{C}_m^{-1}\right)$$

Moreover

$$S(\mathbf{m}) = (\mathbf{m} - \mathbf{m}^{est})^T\; \mathbf{C}_p^{-1}\; (\mathbf{m} - \mathbf{m}^{est})$$

Optimization approach - linear l_2 problem

Assume

$$\mathbf{C}_d = \sigma_d^2 \mathbf{I} \qquad \mathbf{C}_m = \sigma_m^2$$

$$\mathbf{m}^{est} = \mathbf{m}^{apr} + \left(\mathbf{G}^T \mathbf{G} + \frac{\sigma_d^2}{\sigma_m^2} \mathbf{I} \right)^{-1} \mathbf{G}^T \cdot \left(\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr} \right)$$

$$\mathbf{m}^{est} = \mathbf{m}^{apr} + (\mathbf{G}^T \mathbf{G} + \lambda \mathbf{I})^{-1} \mathbf{G}^T \cdot \left(\mathbf{d}^{obs} - \mathbf{G} \cdot \mathbf{m}^{apr} \right)$$

Optimization approach - linear l_2 problem

- Algebraic approach $\equiv$ Optimization approach for linear problem with l_2 norm
- Inversion error: covariance matrix $\mathbf{C}_p$
- Resolution matrix concept:

$$\mathbf{R} \;=\; \mathbf{C}_p^{-1}\,\mathbf{G}^T\mathbf{C}_d^{-1}\cdot\mathbf{G}$$

- The method is NOT robust - large sensitivity to outliers in data

End