Inverse Problems & Emission (Parameter) Estimation

Given a model, a observation, & some understanding of a reasonable parameter value, find the "best parameter"

ex: conc. profile ~~from~~ sat. spectra, RT model
· CO_2 fluxes from trace-gas meas, CTM

"Best parameter"

- ≠ ones that lead to best fit to data
- must account for the fact that the problem is ill-posed (Hadamard, 1901)

Well-posed
- solution exists,
- solution is unique
- solution depends continuously on the data i.e. is stable w.r.t. noise in data

```
clear all
clc
H = [0.16, 0.10; 0.1599, 0.10];
inv(H)
x_true = [1, 1]'
y = H * x_true
x_est1 = inv(H) * y
n = [-0.004, 0.001]';
dy = y + n;
<stop, look at y, dy>
```

$Hx = y$
$\Rightarrow x = H^{-1} y$

```
x_est2 = inv(H) * dy
norm(x_est2 - x_est)/norm(x_est
norm(dy-y)/norm(y)
```

$(\quad) \le \mathrm{cond}(H) \, (\quad)$

$\Rightarrow$ def. of cond #

Multiple LS Regression

$\underset{\sim}{y}$ = data vector $(m \times 1)$

$\underset{\wedge}{x}$ = parameters $(n \times 1)$

$\underset{\approx}{H}$ = linearized model $(m \times n)$

$\hat{y}$ = model est. of data

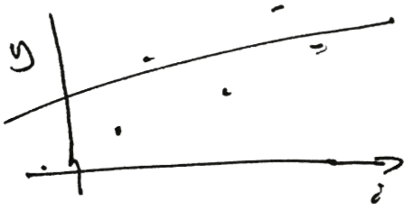
$$Hx = \hat{y}$$

Consider $J(x) = \frac{1}{2} \sum_{i=1}^{\overset{\text{\# data points}}{m}} \left[(Hx)_i - y_i \right]^2 = \frac{1}{2}(Hx-y)^T(Hx-y)$

minimize $J(x)$ $\Rightarrow$ "least squares" Gauss 1795

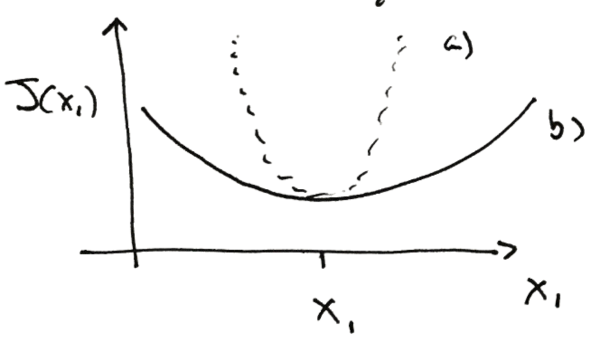
$$\frac{\partial J}{\partial \underset{\sim}{x}} = \frac{1}{2}\left[2\, H^T H x - 2 H^T y \right] = 0$$

$$\Rightarrow x^a = (H^T H)^{-1} H^T y \qquad \text{"normal eqs."}$$



slope: $x_2 \pm$?
int.: $x_1 \pm$?

What is uncertainty in $\underset{\sim}{x}$?



In which case is x_1 more uncertain? $\Rightarrow$ b.

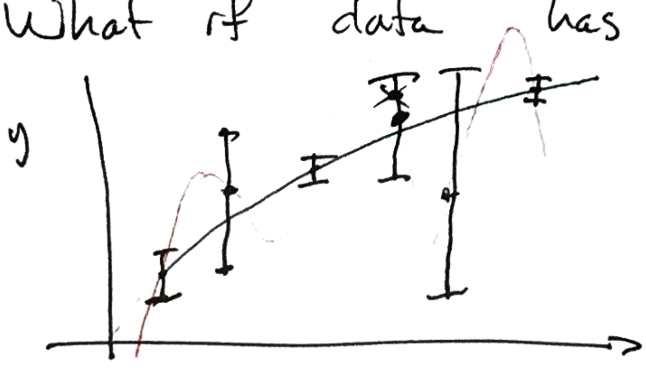
Uncertainty $\propto$ (curvature of J at min)$^{-1}$

$$\propto \left(\frac{\partial^2 J}{\partial x_1^2} \right)^{-1}$$

$\frac{\partial^2 J}{\partial x^2}$ = Hessian $\Rightarrow$ Uncertainty, $P = (\text{Hessian})^{-1}$

$$= (H^T H)^{-1}$$

What if data has errors?



Should fit all points equally? — no!

Weighted L.S.: $J = \frac{1}{2} \sum_{i=1}^{m} \left(\frac{\hat{y}_i - y_i}{\sigma_i} \right)^2$

σ_i : std. dev. of obs i (uncorrelated).

If errors are correlated? $J = \frac{1}{2}(Hx-y)^T R^{-1}(Hx-y)$

where $R \equiv$ obs. error cov. matrix

$$x^a_{LS} = (H^T R^{-1} H)^{-1} H^T R^{-1} y$$ (from $\partial J/\partial x = 0$)

~~(if R ≠ [σ_r σ ...], H doesn't matter (yet))~~

$$P = (H^T R^{-1} H)^{-1}$$

(from $(\partial^2 J/\partial x^2)^{-1}$, or from $cov(x) = A\,cov(y)\,A^T$ for $x = Ay$)

note: for $R = \sigma_r^2 I$, $x_{LS} = (H^T H)^{-1} H^T y$

$P = \sigma_r^2 (H^T H)^{-1}$

<show impulse examp.m, Fig 1 & 2>

{x-perf = (H'*H)^(-1)*H'*y-perf;

note: bad numerical technique

<show Fig 3,4>

Add noise (constant σ_r)

<show Fig 5>

Repeat LS

<show Fig 6> Good solution?

<show Fig 7> !??!

What is happening? Why explode?

{cond(H'*H)

- $(H^T R^{-1} H)^{-1}$ becomes very ill cond., & may not exist
- There are components of $\underline{x}$ that lie in the null space of H, i.e. $H\hat{x} = 0$ for $\hat{x} \neq 0$

If these exist: $H(x + \alpha\hat{x}) = Hx$

so can add $\pm\alpha\hat{x}$ to x_true & it will fit the data just as accurately.

Regularization of ill-posed problems

Restrict problem to find only <u>reasonable</u> values of x.

- Truncated SVD: filter out any component of $x \in N(H)$
- CGLS: solve LS iteratively & halt iterations before we over fit data
- Tikhonov: directly add constraint on $\underset{\sim}{x}$ in $J(\underset{\sim}{x})$
- Bayesian: use prior knowledge of $\underset{\sim}{x}$, $\underset{\sim}{x}^0$.
- Aggregation: use x that is ~~...~~ resolvable

Tikhonov

$$\min J(x) = \frac{1}{2}(Hx-y)^T R^{-1}(Hx-y) + \alpha^2 \underbrace{\|x\|_2^2}_{\text{constraint} \Rightarrow \text{smallest } \underset{\sim}{x}}$$

$$x = (H^T R^{-1} H + \alpha^2 I)^{-1} H^T R^{-1} y$$

stabalize the inverse! Solution is now biased, but has smaller variance.

Other forms

$$J = \ldots + \alpha^2 \|Lx\|_2^2 \quad \text{where } L = \underset{0^{th}}{I},\ \underset{1^{st}}{D},\ \underset{2^{nd}\text{ order}}{D^2}, \ldots$$

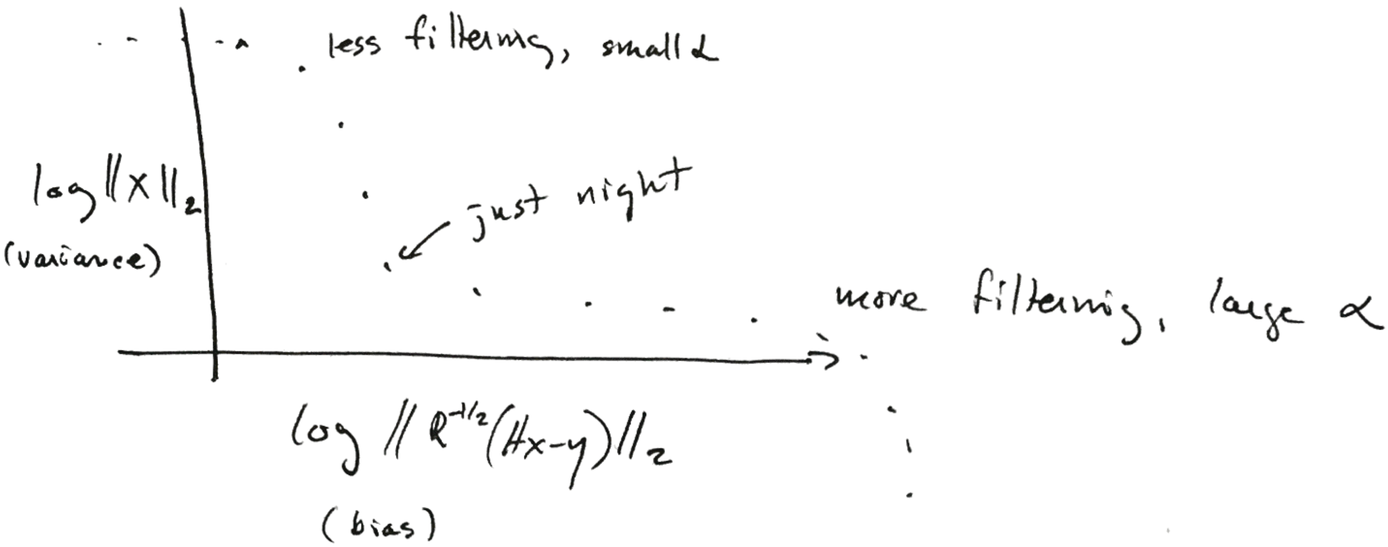
How find α?

- L-curve
- cross validation
- discrepancy principle (χ^2 test).

L-curve

- pick range of α spanning several ~~org~~ orders of magnitude
- find x_α for each

plot $\log \| R^{-1/2}(Hx-y)\|_2$ vs $\log \|x_\alpha\|_2$



Bayesian

$$J = \tfrac{1}{2}(Hx-y)^T R^{-1}(Hx-y) + \tfrac{1}{2}(x-x^b)^T B^{-1}(x-x^b)$$

'background

error in x^b cov.

TSVD

$H = USV^T$ $\qquad [U, S, V] = \text{svd}(H)$ in MATLAB

U = $m \times m$ orthogonal matrix
Columns $\underset{\sim}{u}_i$ span data space $\mathbb{R}^m$
left singular vectors

V = $n \times n$ matrix orthogonal
Columns $\underset{\sim}{v}_i$ span model space
param

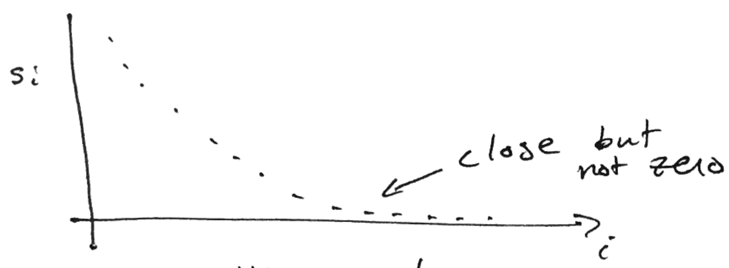
S = $m \times n$ diag matrix, non-neg. elements s_i,
ordered $s_1 > s_2 > \ldots$
<show SVD of H>

$p = \text{rank}(H) = \#$ of non-zero s_i

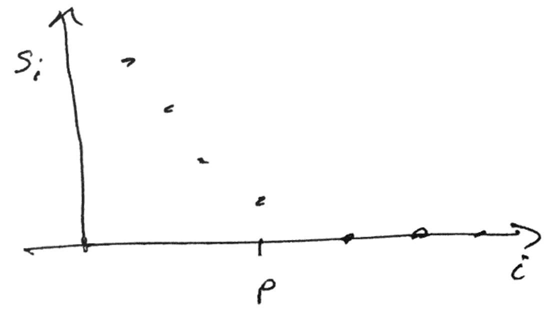
$$S = \begin{bmatrix} S_p & 0 \\ 0 & 0 \end{bmatrix}$$

$$H = [U_p \;\; U_0] \begin{bmatrix} S_p & 0 \\ 0 & 0 \end{bmatrix} [V_p \;\; V_0]^T$$

Singular Values



ill-cond.



rank deficient

$$\hat{x}_{LS} = (H^T H)^{-1} H^T y$$

$$= \left[(USV^T)^T USV^T\right]^{-1} (USV^T)^T y$$

use: $(AB)^T = B^T A^T$

orth. $\begin{cases} U^T U = I \\ V^T V = I \end{cases}$

$$= \left[V S U^T U S V^T\right]^{-1} V S U^T y$$

$$= \left[V S^2 V^T\right]^{-1} V S U^T y$$

$$= V S^{-2} V^T V S U^T y$$

$$= V S^{-1} U^T y$$

$$= \sum_{i=1}^{n} \left(\frac{u_i^T y}{s_i}\right) v_i$$ is linear comb. of v_i's.

If H is rank deficient: $s_i = 0$ for $i > p$

i.e. $n > p$ (fat H)

$m > p$ (skinny H)

$\hat{x}_{LS}$ would explode, unless $u_i^T y$ is zero, which is unlikely since y is real data w/noise

Avoid?

$$x_p \equiv \sum_{i=1}^{p} \left(\frac{u_i^T y}{s_i}\right) v_i$$

note: set $\frac{1}{s_{i>p}} = 0$, i.e. $\frac{1}{0} = 0$? odd!

$$= V_p S_p^{-1} U_p^T y$$

where V_p is 1st p columns of V

i.e., $S = \begin{bmatrix} S_p & 0 \\ 0 & 0 \end{bmatrix}$

$null(H) = span\{V_{p+1}, \ldots, V_N\}$

$H^+ \equiv V_p S_p^{-1} U_p^T$ is Moore-Penrose pseudo inverse, always exists

"pinv" in MATLAB

ill-~~cond~~ posed: s_i ~~set~~ decay smoothly towards zero
(the quicker → more ill-posed)

truncate series at $p = p'$

$$x_{p'} = \sum_{i=1}^{p'} \frac{u_i^T y}{s_i} v_i = V_{p'} S_{p'}^{-1} U_{p'}^T y$$

How do we know where to truncate?

Consider discrepancy principle:

- fit each obs to within $\pm \sigma_r$
- $\sum \left[\frac{(H\cancel{x})_i - y_i}{\sigma_r} \right]^2$ should $\approx m$
- or $\frac{\| Hx - y \|_2^2}{m} \approx \sigma_r^2$

So, check different p' $(1 \ldots \max(n, m))$, recalculate $\frac{\| Hx - y \|_2^2}{m}$, stop when $\approx \sigma_r^2$.