Underdetermined systems and Adjoints

M = number of equations

N = number of param.

$$\underline{y} = \underline{\underline{E}}\,\underline{x} + \underline{n}$$

$\underline{\underline{E}}_{M \times N}$

we have seen that using LSQ

$$J = \underbrace{\underline{n}^T \underline{\underline{R}}_{nn}^{-1} \underline{n}} + \underline{x}^T \underline{\underline{C}}_{xx}^{-1} \underline{x}$$

minimize observation-data misfit within observation error $\underline{\underline{R}}_{nn}$

Adjust model parameters with the typical size of $\underline{x}$ given by $\underline{\underline{C}}_{xx}$

solutions were found

$$\left\{\begin{aligned}
\hat{\underline{x}} &= \left(\underline{\underline{E}}^T \underline{\underline{R}}_{nn}^{-1} \underline{\underline{E}} + \underline{\underline{C}}_{xx}^{-1}\right)^{-1} \underline{\underline{E}}^T \underline{\underline{R}}_{nn}^{-1} \underline{y} \\
\hat{\underline{\underline{C}}}_{xx} &= \hat{\underline{\underline{E}}} \underline{\underline{R}}_{nn} \hat{\underline{\underline{E}}} \qquad \hat{\underline{\underline{E}}} = \left(\underline{\underline{E}}^T \underline{\underline{R}}^{-1} \underline{\underline{E}} + \underline{\underline{C}}_{xx}^{-1}\right) \underline{\underline{E}}^T \\
\hat{\underline{n}} &= \underline{y} - \underline{\underline{E}}\,\hat{\underline{x}}
\end{aligned}\right.$$

this is the general form.

using the matrix inversion lemma

$$\underbrace{\left(\underline{\underline{E}}^T \underline{\underline{R}}_{nn}^{-1} \underline{\underline{E}} + \underline{\underline{C}}_{xx}^{-1}\right)}_{N \times N} \underline{\underline{E}}^T \underline{\underline{R}}_{nn}^{-1} = \underline{\underline{C}}_{xx} \underline{\underline{E}}^T \underbrace{\left(\underline{\underline{E}}\,\underline{\underline{C}}_{xx} \underline{\underline{E}}^T + \underline{\underline{R}}_{nn}\right)^{-1}}_{M \times M}$$

$$\left\{\ \hat{x} = \underline{\underline{C}}_{xx} \underline{\underline{E}}^T \left(\underline{\underline{E}}\,\underline{\underline{C}}_{xx} \underline{\underline{E}}^T + \underline{\underline{R}}_{nn}\right)^{-1} \underline{y}\right.$$

Now assume the case $M < N$ the underdetermine problem.

Assume you have a linear system that has a linear constraint you want to satisfy exactly

$$\underline{y} = \underline{\underline{E}}\,\underline{x}$$

$$\begin{cases} x_1 + x_2 + x_3 = 1 \\ x_1 - x_2 + x_3 = 2 \end{cases}$$

example system

you want to find a solution so that $x_1^2 + x_2^2 + x_3^2$ is small. In other words you know a priori something about the stats. of your unknowns.

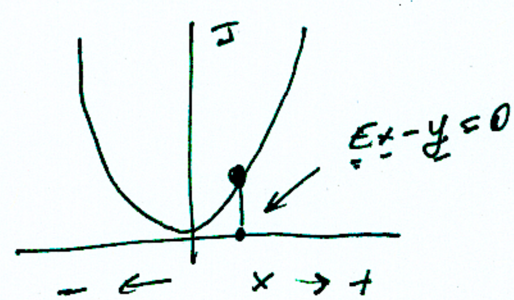
We can define

$$J = \underline{x}^T \underline{\underline{C}}_{xx}^{-1}\,\underline{x} + 2\underline{\mu}^T(\underline{\underline{E}}\,\underline{x} - \underline{y})$$

bound on $\underline{x}$

enforce the linear constraint exactly by adding a lagrange multiplier $\underline{\mu}$ term.

$$\begin{cases} \frac{1}{2}\frac{\partial J}{\partial \underline{x}} = 0 = \underline{\underline{C}}^{-1}\underline{x} - \underline{\underline{E}}^T\underline{\mu} \\ \frac{1}{2}\frac{\partial J}{\partial \underline{\mu}} = 0 = \underline{\underline{E}}\,\underline{x} - \underline{y} \end{cases}$$



1. $\hat{\underline{x}} = \underline{\underline{C}}\,\underline{\underline{E}}^T\underline{\mu}$

2. $\underline{\underline{E}}\,\underline{x} - \underline{y} = \underline{\underline{E}}\,\underline{\underline{C}}\,\underline{\underline{E}}^T\underline{\mu} - \underline{y} \implies \underline{\mu} = (\underline{\underline{E}}\,\underline{\underline{C}}\,\underline{\underline{E}}^T)^{-1}\underline{y}$

$$\hat{\underline{x}} = \underline{\underline{C}}\,\underline{\underline{E}}^T(\underline{\underline{E}}\,\underline{\underline{C}}\,\underline{\underline{E}}^T)^{-1}\underline{y}$$

finally

$$\begin{cases} \hat{\underline{x}} = \underline{\underline{C}}\,\underline{\underline{E}}^T(\underline{\underline{E}}\,\underline{\underline{C}}\,\underline{\underline{E}}^T)^{-1}\underline{y} \\ \hat{\underline{n}} \equiv 0 \\ \hat{\underline{\underline{C}}}_{xx} = \phi \quad \text{because it is a function of } \underline{\underline{R}}_{nn} \equiv 0 \end{cases}$$

this is equivalent to previous LSQ estimate with $\underline{\underline{R}}_{nn}=0$ (using inversion lemma)

If we had account in J for error $\underline{n}$

$$J = \underline{x}^T\underline{\underline{C}}^{-1}\underline{x} + \underline{n}^T\underline{\underline{R}}_{nn}\underline{n} - 2\underline{\mu}^T(\underline{\underline{E}}\,\underline{x}+\underline{n}-\underline{y})$$

we retrieve LSQ fully!

$\begin{cases} \underline{n} = 0 & \text{STRONG CONSTRAINT} \\ \underline{n} \neq 0 & \text{WEAK CONSTRAINT} \end{cases}$

Adjoints: interpretations and uses

often we have seen the operator $\underline{\underline{E}}^T$.

such operator is called the "Adjoint" of $\underline{\underline{E}}$

Example 1: $y = \underline{\underline{E}}\,\underline{x} = [1\ \tfrac{1}{2}\ 1]\begin{bmatrix}1\\ 2\\ 1\end{bmatrix} = 3$ scalar

$\underline{r} = \underline{\underline{E}}^T y = \begin{bmatrix}1\\ 1/2\\ 1\end{bmatrix}[3] = \begin{bmatrix}3\\ 3/2\\ 3\end{bmatrix}$ I can reconstruct the weightings

Assume a more interesting example

$\underline{\underline{E}}$ is the forward integral model of eq.

$$\cancel{\frac{dy}{dt}} \qquad \frac{dy}{dt} = \alpha y$$

centered scheme

$$\frac{y_{n+1} - y_{n-1}}{2\Delta t} = \alpha y_n$$

and assume we run $n = 0 \ldots 3$

Euler scheme

$$\frac{y_{n+1} - y_n}{\Delta t} = \alpha y_n$$

$n = 0$ (Euler step)

$$y_{n+1} = y_{n-1} + \alpha 2\Delta t\, y_n$$

$$y_{n+1} = y_n + \alpha \Delta t\, y_n = (1 + \alpha \Delta t)\, y_n$$

~~Using Euler we find that y_{n+3}~~

$n = 0$ $\quad \dfrac{y_1 - y_0}{\Delta t} = \alpha y_0 \longrightarrow y_1 = (1 + \alpha \Delta t) y_0$

$$\frac{y_2 - y_0}{2\Delta t} = \alpha y_1 \longrightarrow y_2 = y_0 + \alpha 2\Delta t\, y_1$$

$$\frac{y_3 - y_1}{2\Delta t} = \alpha y_2 \longrightarrow y_3 = y_1 + \alpha 2\Delta t\, y_2$$

We can write the system in Matrix

$$\underbrace{\begin{bmatrix} y_1 \\ y_2 \\ \hat{y}_3 \end{bmatrix}}_{\underline{y}} = \underbrace{\begin{bmatrix} 1+\alpha\Delta t & 0 & 0 \\ 1 & \alpha 2\Delta t & 0 \\ 0 & 1 & \alpha 2\Delta t \end{bmatrix}}_{\underline{\underline{E}}} \underbrace{\begin{bmatrix} y_0 \\ y_1 \\ y_2 \end{bmatrix}}_{\underline{x}}$$

$\underline{\underline{E}}$ the forward model $y_0 \rightarrow \hat{y}_3$

now let us $\underline{\underline{E}}^T \underline{y}$

$$\underbrace{\begin{bmatrix} 1+\alpha\Delta t & 1 & 0 \\ 0 & \alpha 2\Delta t & 1 \\ 0 & 0 & \alpha 2\Delta t \end{bmatrix}}_{\underline{\underline{E}}^T} \underbrace{\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}}_{\underline{y}} \overset{?}{=} \underbrace{\begin{bmatrix} \hat{y}_0 \\ y_1 \\ y_2 \end{bmatrix}}_{\underline{x}}$$

$$y_0 = y_1(1+\alpha\Delta t) + y_2$$

$$y_1 = 2\alpha\Delta t\, y_2 + y_3 \longleftarrow -\left(\frac{y_3 - y_1}{2\Delta t}\right) = \alpha y_2$$

$$y_2 = y_3\, \alpha\Delta t$$

$$\parallel$$

$$-\frac{dy}{dt} = \alpha y$$ backward model.

(propagation)

How to use this "backward" information?

"SENSITIVITY ANALYSIS"

Assume 4 our ADV-DIFF problem of our pollutant

$$\underline{y} = \underline{\underline{E}}\,\underline{x} + \underline{n}$$

$$\underline{x} = \underline{T}(t=0) \qquad \underline{y} = T(t=100)$$

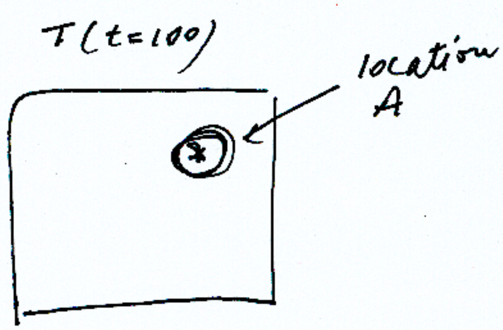
$\underline{\underline{E}}$ is the integral solution that carries

$$T(t=0) \longrightarrow T(t=100)$$

$$\underline{T}(t_N) = \underline{\underline{R}}(t_0, t_N)\,\underline{T}(t_0) \qquad \begin{array}{l} t_0 = 0 \\ t_N = 100 \end{array}$$

forward propagator dynamics $\equiv \underline{\underline{E}}$ in our case.

<u>Question</u>: what is the variance at location A sensitive to? In other words where do I change my initial conditions to maximize the variance at A?



Let us define a cost function

$$J = \underline{y}^T \underline{\underline{W}}\, \underline{y} \quad \text{units of } ppm^2 \qquad \underline{\underline{W}} = \underline{w}\,\underline{w}^T$$

$$\underline{w} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ \vdots \end{bmatrix} \leftarrow \text{location } A$$

so now we want to know how J changes with respect to $\underline{x}$.

$$\frac{\partial J}{\partial \underline{x}} = \left((\underline{\underline{E}}x)^T \underline{\underline{W}}\, \underline{\underline{E}} \right)^T = \underline{\underline{E}}^T \underline{\underline{W}} (\underline{\underline{E}}\underline{x}) = \underbrace{\underline{\underline{E}}^T}_{\text{adjoint model}} \underline{\underline{W}}\, \underline{y}$$

this is called the gradient vector or adjoint solution.

NOTE: The units are $\dfrac{ppm^2}{\text{over } x}$