Regularized Least-Squares and Gauss-Newton Methods

Gabriel Hugh Elkaim

Regularized Least-Squares and Gauss-Newton Methods

- Multi-objective Least-Squares
- Regularized Least-Squares
- Non-linear Least Squares
- Gauss-Newton Method

$\boxed{\text{TANSTAAFL}}$

Multi-objective Least-Squares (1.3)

$J_1 = \|A x - y\|^2$ small

$J_2 = \|F x - g\|^2$ small.

$x \in \mathbb{R}^n \leftarrow$ some x to min. both J_1 & J_2 simultaneously

objectives are competing

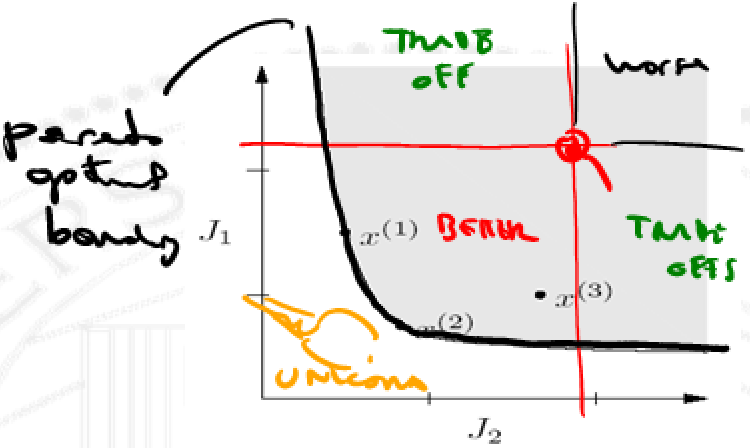
can only make J_1 small at the cost of increase in J_2

$F = I \quad g = 0$ min $\|Ax - y\|$ subj min $\|x\|$.

Multi-objective Least-Squares (2.3)

plot $J_1(x)$ and $J_2(x)$

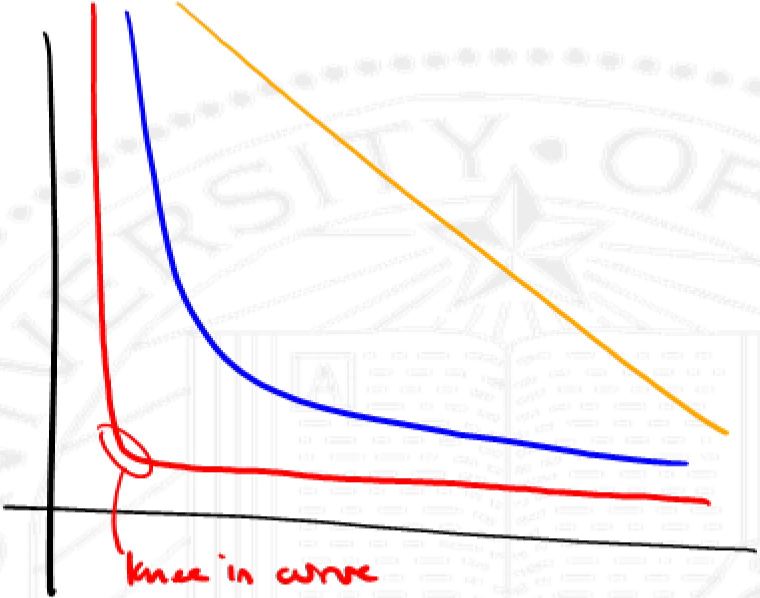
efficient boundary.



Multi-objective Least-Squares (3.3)

$J_1 = \|Ax - y\|^2$

$J_2 = \|Fx - g\|^2$



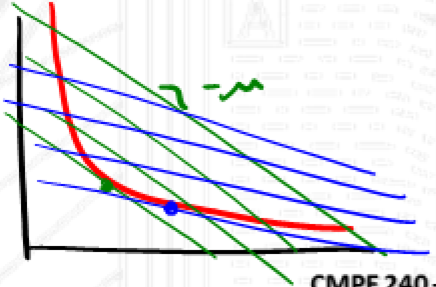
Weighted-sum Objective (1.3)

$$J = J_1 + \mu J_2 = \|Ax - g\|^2 + \mu \|Fx - g\|^2$$

$\mu \geq 0$ gives the relative weight between J_1 & J_2

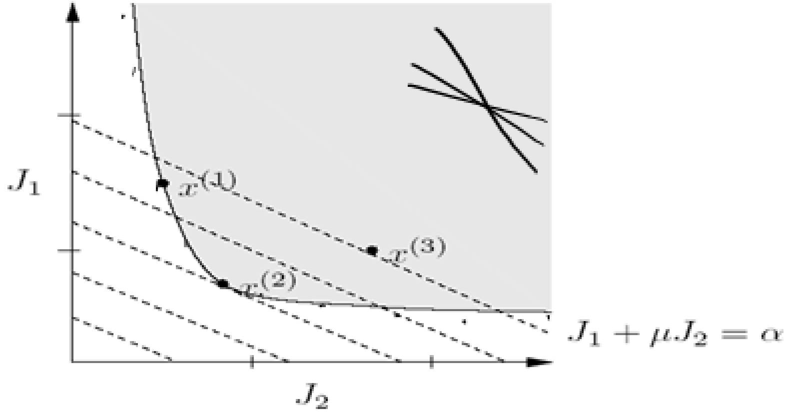
↑ kind of exchange rate

$J_1 + \mu J_2 = \alpha$



Weighted-sum Objective (2.3)

Vary μ from $0 \to \infty$



Minimizing Weighted-Sum Objective

$$\|Ax-y\|^2 + \mu\|Fx-g\|^2 = \left\| \underbrace{\begin{bmatrix} A \\ \sqrt{\mu}F \end{bmatrix}}_{\tilde{A}} x - \underbrace{\begin{bmatrix} y \\ \sqrt{\mu}\, g \end{bmatrix}}_{\tilde{y}} \right\|^2$$

$$\min_x \|\tilde{A}x - \tilde{y}\|^2 \longrightarrow x_{ls} = (\tilde{A}^T\tilde{A})^{-1}\tilde{A}^T\tilde{y}$$

$$\boxed{x_{ls} = (A^TA + \mu F^TF)^{-1}(A^Ty + \mu F^Tg)}$$

$$\mu = 0 \longrightarrow x_{ls} = (A^TA)^{-1}A^Ty \qquad x_{ls} \rightarrow Ax = y$$

$$\mu = \infty \longrightarrow x_{ls} = (F^TF)^{-1}F^Tg \qquad x_{ls} \rightarrow Fx = g$$

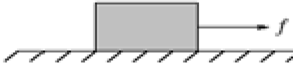
Minimizing Weighted-Sum Objective

$$\min_x \|Ax - y\|^2 + \mu\|Dx\|^2 + \lambda\|x\|^2$$

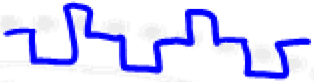
$$D = \begin{bmatrix} 1 & -1 & & & 0 \\ & \ddots & \ddots & & \\ & & \ddots & \ddots & \\ 0 & & & & \end{bmatrix}$$

"smoothness"

Example: Mass-Force Program



$i = 1..10$



$y \in \mathbb{R}$ position @ $t = 10$ $y = a^T x$ $a \in \mathbb{R}^{10 \times 1}$

$J_1 = (y-1)^2$ final position error squared.

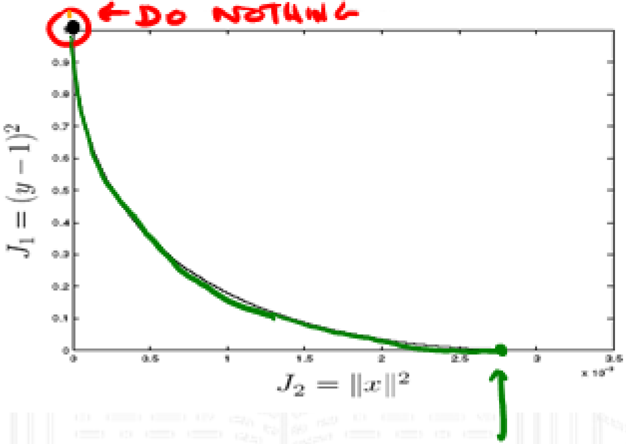
$J_2 = \|x\|^2$ sum square of forces

$$J = (a^T x - 1)^2 + \mu x^T x$$

$$\boxed{x_{opt} = (a a^T + \mu I)^{-1} a \cdot 1.}$$

Example: Mass-Force Program

BEING PERVERSE



Regularized Least-Squares (1.3)

when $F = I$ $g = \phi$. ← "ridge regression"

$J_1 = \|Ax - y\|^2$ $J_2 = \|x\|^2$

$$\boxed{x_{ls} = (A^TA + \mu I)^{-1} A^T y}$$

Regularized least squares solution $Ax \approx y$.

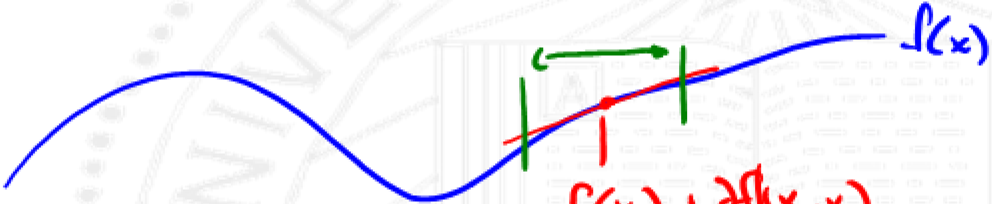
Tychonov Regularization

$\mu > 0$ works for <u>any</u> A.

Regularized Least-Squares (2.3)

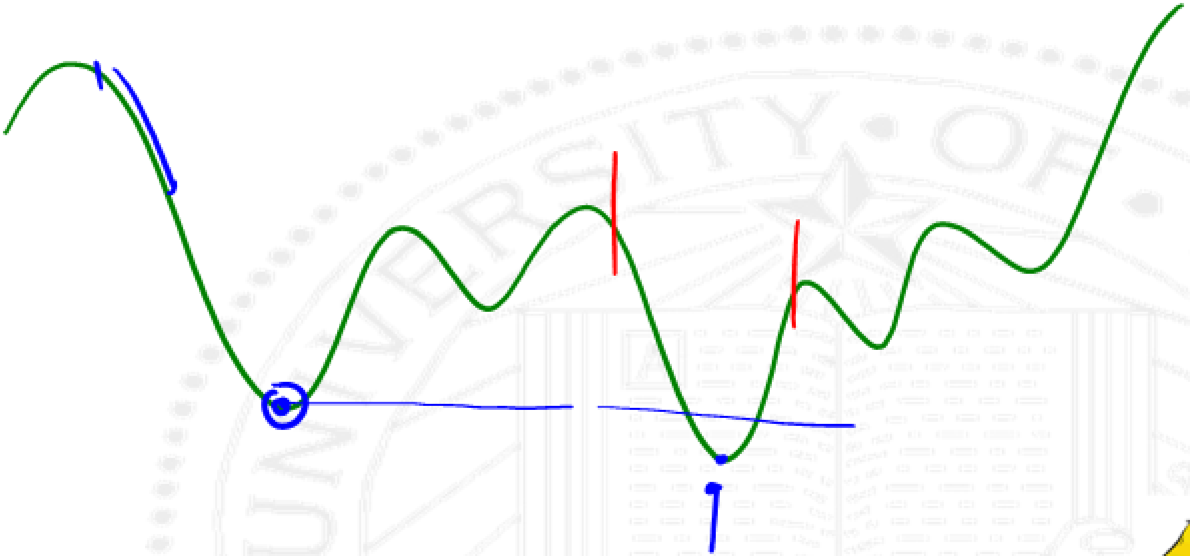
$Ax - y$ sensor residual.

prior information: x is small.



$$f(x_0) + \left.\frac{\partial f}{\partial x}\right|_{x=x_0}(x - x_0)$$

Regularized Least-Squares (3.3)



Non-Linear Least Squares (1.3)

find $x \in \mathbb{R}^n$ $\qquad \min_x \|r(x)\|^2 = \sum_{i=1}^{m} r_i(x)^2$

$r : \mathbb{R}^n \to \mathbb{R}^m$

In general, very hard to solve exactly

Reduces to (linear) LS if $r(x) = Ax - y$.

lots of heuristics to compute a locally optimal solution.

Non-Linear Least Squares (2.3)

1. Guess $x \leftarrow x_0$ initial condition
2. linearize r near current x. { calculus, perturbations, Partial differentials }
3. solve linear least squares on 2
4. Update x.

$$r(x) \simeq Ax - b \text{ near } x_0.$$

Non-Linear Least Squares (3.3)

① linearize r near around $x^{(k)}$

$$r(x) \simeq r(x^{(k)}) + Dr(x^{(k)}) \big[\overbrace{x - x^{(k)}}^{\Delta x} \big]$$

$$Dr \triangleq \text{Jacobian} \rightarrow Dr_{ij} = \frac{\partial r_i}{\partial x_j} \quad m \times n.$$

$$r(x) \simeq r(x^{(k)}) + Dr(x^{(k)})\, \Delta x \simeq A^{(k)} \Delta x - b^{(k)}$$

$$\boxed{\begin{aligned} A^{(k)} &= Dr(x^{(k)}) \\ b^{(k)} &= Dr(x^{(k)}) - r(x^{(k)}) \end{aligned}}$$

NLLS Example: Position from Ranges

$$\|r(x)\|^2 \approx \|A^{(k)}\delta x - b^{(k)}\|^2$$

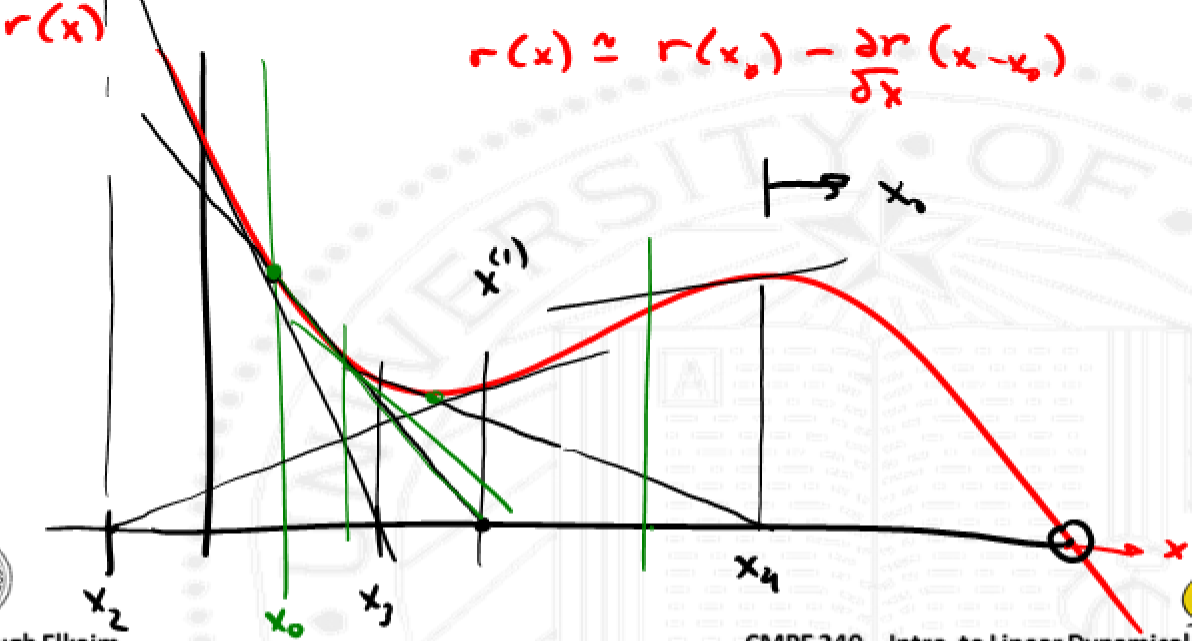
$$x^{(k+1)} = \underbrace{\left(A^{(k)T}A^{(k)}\right)^{-1} A^{(k)T} b^{(k)}}_{\delta x_{ls}} + x^{(k)}$$

$$J = \|A^{(k)}\delta x - b^{(k)}\|^2 + \mu \underbrace{\|x - x^{(k)}\|^2}_{\text{TRUST REGION}}$$

$F = I \quad g = 0.$

Gauss-Newton method for NLLS (1.3)

$$r(x) \simeq r(x_0) - \frac{\partial r}{\delta x}(x - x_0)$$



Gauss-Newton method for NLLS (2.3)

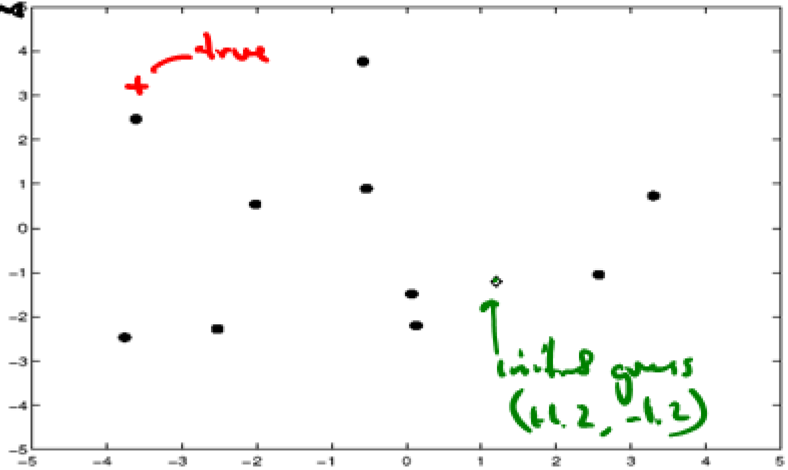
Gauss-Newton method for NLLS (3.3)

Gauss-Newton: Example (1.4)

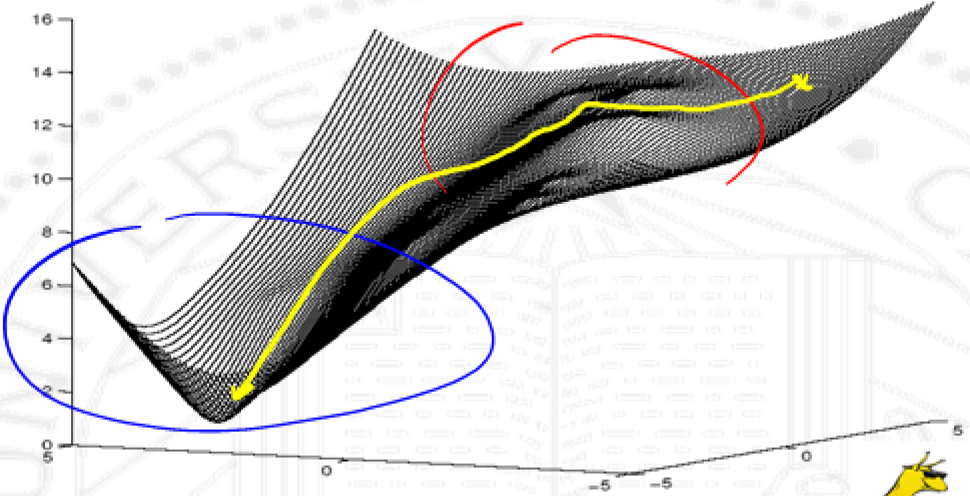
10 beacons @ known points

(−3.6, 3.2)

range ~ N(0, .5m)



Gauss-Newton: Example (2.4)

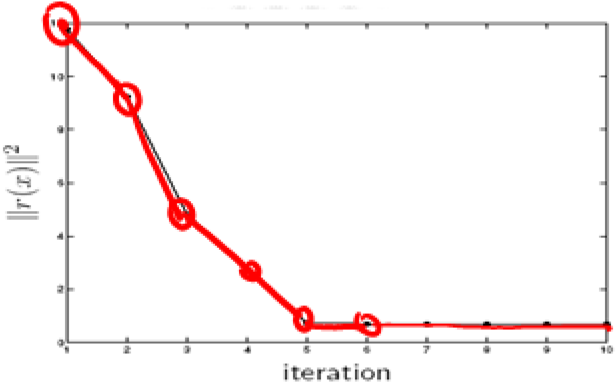


Gauss-Newton: Example (3.4)

converge to the global min

$\hat{x} = (-3.3, 3.3)$

$\|x - \hat{x}\|^2 = 0.31$

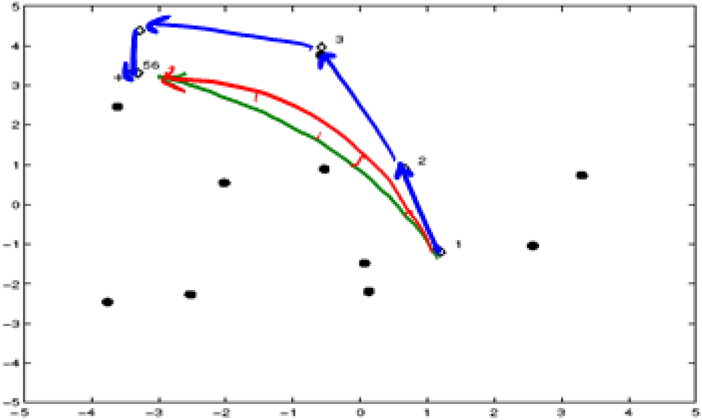


Gauss-Newton: Example (4.4)

use trust region

$\Delta x: \quad x^{+} := x^{-} + \dfrac{\Delta x}{10}$

$\|A^{(k)}\Delta x - b^{(k)}\|^2 + \mu\|\Delta x\|^2$



Variation on Gauss-Newton

Armero Steps
Simulated Annealing

$\Bigg\}$ vary step size.

$$\frac{\delta x}{\mu}$$

Questions?