

Multi-objective least-squares

Stephen Boyd and Sanjay Lall

EE263
Stanford University

Multi-objective least-squares

in many problems we have two (or more) objectives

- ▶ we want $J_1 = \|Ax - y\|^2$ small

- ▶ and also $J_2 = \|Fx - g\|^2$ small

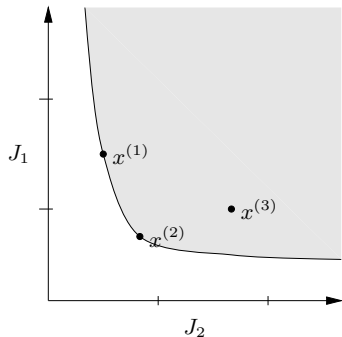
($x \in \mathbb{R}^n$ is the variable)

- ▶ usually the objectives are *competing*

- ▶ we can make one smaller, at the expense of making the other larger

common example: $F = I$, $g = 0$; we want $\|Ax - y\|$ small, with small x

Plot of achievable objective pairs



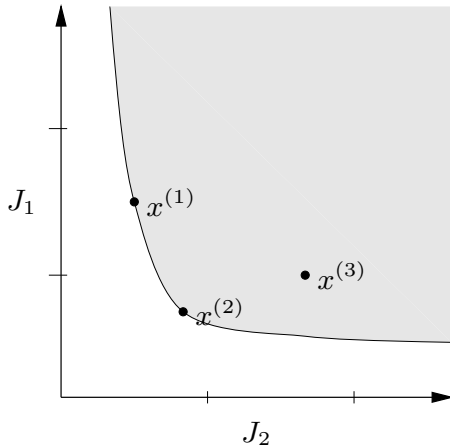
- plot (J_2, J_1) for every x
- $x \in \mathbb{R}^n$, but this plot is in $\mathbb{R}^2$
- point labeled $x^{(1)}$ is really $(J_2(x^{(1)}), J_1(x^{(1)}))$

Optimal trade-off curve

- ▶ shaded area shows (J_2, J_1) achieved by some $x \in \mathbb{R}^n$
- ▶ clear area shows (J_2, J_1) not achieved by any $x \in \mathbb{R}^n$
- ▶ boundary of region is called *optimal trade-off curve*
- ▶ corresponding x are called *Pareto optimal* for the two objectives $\|Ax - y\|^2, \|Fx - g\|^2$

three example choices of x : $x^{(1)}, x^{(2)}, x^{(3)}$

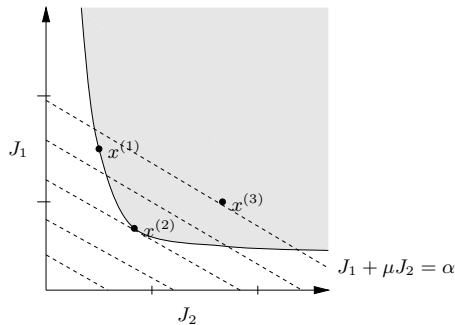
- ▶ $x^{(3)}$ is worse than $x^{(2)}$ on both counts (J_2 and J_1)
- ▶ $x^{(1)}$ is better than $x^{(2)}$ in J_2 , but worse in J_1



Weighted-sum objective

- ▶ to find Pareto optimal points,
i.e., x 's on optimal trade-off curve,
we minimize *weighted-sum objective*

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



- ▶ parameter $\mu \geq 0$ gives relative weight between J_1 and J_2
- ▶ points where weighted sum is constant, $J_1 + \mu J_2 = \alpha$, correspond to line with slope $-\mu$ on (J_2, J_1) plot
- ▶ $x^{(2)}$ minimizes weighted-sum objective for μ shown
- ▶ by varying μ from 0 to $+\infty$, can sweep out entire *optimal tradeoff curve*

Minimizing weighted-sum objective

can express weighted-sum objective as ordinary least-squares objective:

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

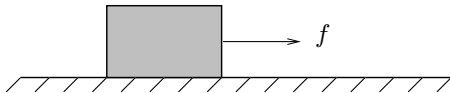
where

$$\tilde{A} = \begin{bmatrix} A \\ \sqrt{\mu}F \end{bmatrix}, \quad \tilde{y} = \begin{bmatrix} y \\ \sqrt{\mu}g \end{bmatrix}$$

hence solution is (assuming $\tilde{A}$ full rank)

$$\begin{aligned}x &= (\tilde{A}^\top \tilde{A})^{-1} \tilde{A}^\top \tilde{y} \\ &= (A^\top A + \mu F^\top F)^{-1} (A^\top y + \mu F^\top g)\end{aligned}$$

Example



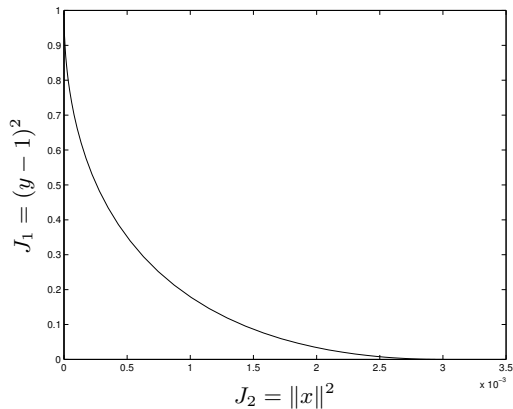
- ▶ unit mass at rest subject to forces x_i for $i - 1 < t \leq i$, $i = 1, \dots, 10$
- ▶ $y \in \mathbb{R}$ is position at $t = 10$; $y = a^\top x$ where $a \in \mathbb{R}^{10}$
- ▶ $J_1 = (y - 1)^2$ (final position error squared)
- ▶ $J_2 = \|x\|^2$ (sum of squares of forces)

weighted-sum objective: $(a^\top x - 1)^2 + \mu \|x\|^2$

optimal x :

$$x = (aa^\top + \mu I)^{-1} a$$

Optimal trade-off curve



- ▶ upper left corner of optimal trade-off curve corresponds to $x = 0$
- ▶ bottom right corresponds to input that yields $y = 1$, *i.e.*, $J_1 = 0$

Regularized least-squares

when $F = I$, $g = 0$ the objectives are

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

minimizer of weighted-sum objective,

$$x = (A^T A + \mu I)^{-1} A^T y,$$

is called *regularized* least-squares (approximate) solution of $Ax \approx y$

- ▶ also called *Tychonov regularization*
- ▶ for $\mu > 0$, works for *any* A (no restrictions on shape, rank ...)

estimation/inversion application:

- ▶ $Ax - y$ is sensor residual
- ▶ prior information: x small
- ▶ regularized solution trades off sensor fit, size of x