March 30, 2026

Announcements

- Exam 3

Goals

- NLSQ
  - CM
- Constrained optimization

  $f: \mathbb{R}^n \to \mathbb{R}$

  Examples from Nocedal & Wright.

NLSQ

$$\vec{r}(\vec{x}) = \vec{y} - \vec{a}(\vec{x})$$

$$\varphi(\vec{x}) = \frac{1}{2} \vec{r}^T(\vec{x})\, \vec{r}(\vec{x})$$

$$H_\varphi \approx J^T J$$

$$J^T J s = -J^T r$$

$$J \vec{s} \approx -r$$

Levenberg-Marquardt

$$(J^T J + \mu_k I_n) s_k = -J^T r$$

If $\mu_k \to \infty$:

$$s_k = \frac{1}{\mu_k}(-J^T r) = -\frac{1}{\mu_k} \nabla \varphi$$

steepest descent

If $\mu_k = 0$: Gauss-Newton

Indicator: $\rho_k = \dfrac{\varphi(x_k) - \varphi(x_k + s_k)}{\text{predicted decrease from derivatives}}$

If $\rho_k \approx 1$: decrease μ_k
$\rho \approx 0$: increase μ_k

"trust region"

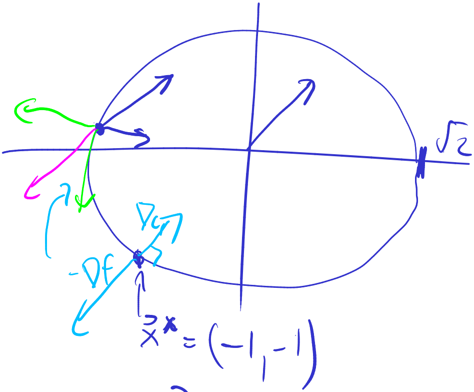
Constrained opt

$$\min f(\vec{x})$$

$$c_i(\vec{x}) = 0 \quad \leftarrow \text{equality}$$

$$c_j(\vec{x}) \geq 0 \quad \leftarrow \text{inequality}$$

$f(x_1, x_2) = x_1 + x_2$ $\qquad$ $c_1(x_1, x_2) = 2 - (x_1^2 + x_2^2) = 0$



$$\nabla f = \partial_i^T \vec{\lambda}$$

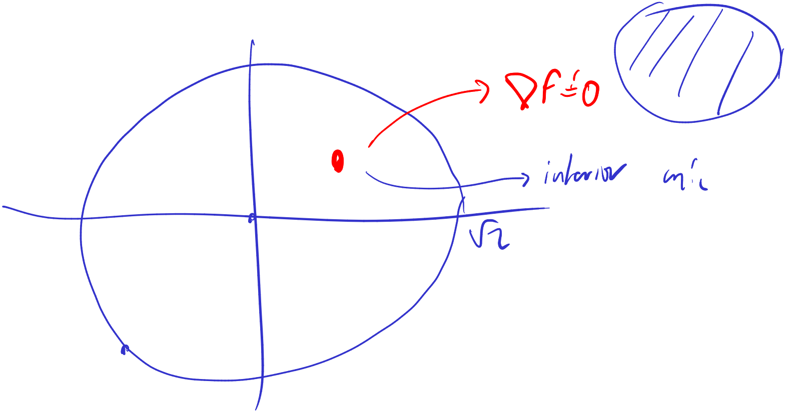
$$\mathcal{L}(\vec{x}, \vec{\lambda}) = f(\vec{x}) - \vec{\lambda}^T \cdot \vec{c}(\vec{x})$$

$$\nabla_\lambda \mathcal{L} \overset{!}{=} 0 \quad \to \quad \vec{c}(\vec{x}) = \vec{0}$$

$$\nabla_{\vec{x}} \mathcal{L} = \nabla f - \lambda^T \partial \vec{c}(\vec{x}) \overset{!}{=} 0$$

$$\nabla \mathcal{L} \begin{pmatrix} x_1 \\ x_1 \\ \vdots \\ x_n \\ \lambda_1 \\ \vdots \\ \lambda_m \end{pmatrix} = \begin{pmatrix} \vdots \\ \vdots \\ c_1(\vec{x}) \\ \\ \end{pmatrix} = 0$$

$$f(x_1, x_2) = x_1 + x_2 \qquad c_1(x_1, x_2) = 2 - (x_1^2 + x_2^2) \geq 0$$



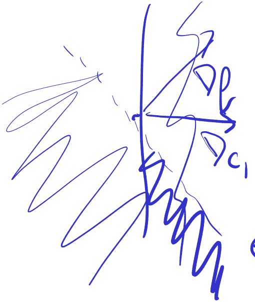
Feasible descent dir $\vec{s}$:

$$A = \{ \nabla f(\vec{x})^T \vec{s} < 0 \}$$

$$B = \{ \nabla c_1(\vec{x})^T s \geq 0 \}$$

necessary cond for feasible min: no descent dir. (dir.)

$$A \cap B = \emptyset$$



← cone of feasible descent dir.

empty if $\nabla f = \lambda \nabla c_1$, same as before

force $\lambda \geq 0$

$$c_1(x) \cdot \lambda_1(x) = 0$$

Necessary condi:

KKT conditions

$$\nabla \mathcal{L} \stackrel{!}{=} 0$$

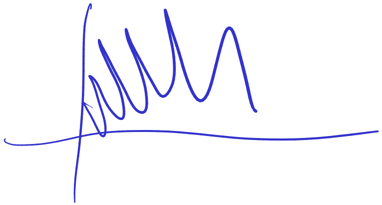
constraints

$$\vec{\lambda} \geq \vec{0}$$

$$\vec{\lambda}^T \vec{c} = 0 \quad \text{(complementarity)}$$

Cones

$K \subseteq \mathbb{R}^n$ $\qquad$ $K = \{ B\vec{y} + C\vec{w} \, : \, \vec{y} \geq 0 \}$



Farkas' lemma

$$K = \{ B\vec{y} + C\vec{w} : \vec{y} \geq \vec{0} \}$$

Given any vector $g \in \mathbb{R}^n$, either

- $\vec{g} \in K$, or
- there exists a $\vec{d}$:

$$\vec{g}^T \vec{d} < 0, \quad \vec{B}^T \vec{d} \geq 0, \quad C^T \vec{d} = 0 \}$$

$$K = \left\{ \sum_i \lambda_i \nabla c_i(\vec{x}^*) \;:\; \lambda_i \geq 0 \right\}$$

$$g = \nabla f(\vec{x}^*)$$