2/19: KKT Conditions & Lagrangians

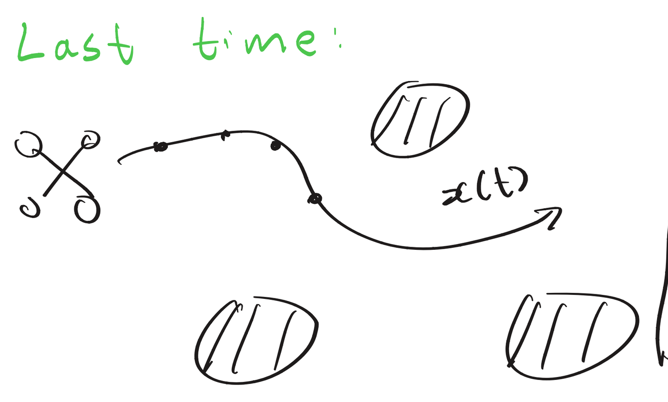
Announcements:

- Project proposals due <u>today</u>, 2/19
  - groups of ≤ 3
  - 1 pg. max, excluding ref.
- HW2 out, due 3/5

Today:

- constraints via projection
- Lagrange multipliers
- KKT
- Augmented Lagrangian methods

Last time:



$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } h_i(x) = 0 \quad i = 1, \dots, \ell$$
$$g_j(x) \leq 0 \quad j = 1, \dots, m$$

constrained optimization

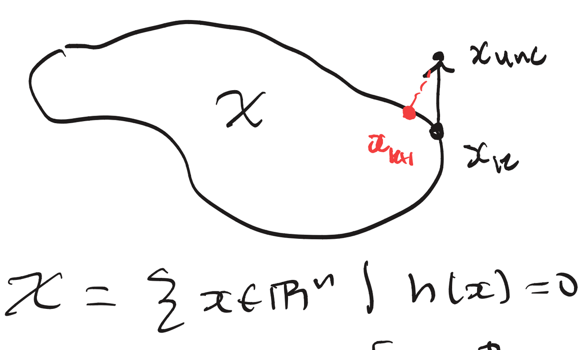
<u>Option 1:</u> Reformulate

Find $T(\hat{x})$ where $h_j(T(\hat{x})) = 0$, $g_j(T(\hat{x})) \leq 0$

Then we can solve an unconstrained problem

$$\min_{\hat{x}} f(T(\hat{x}))$$

Option 2: Project



Key idea: reproject onto $\mathcal{X}$

$x_{k+1} = \Pi_{\mathcal{X}}(x_{unc})$ ← unconstrained step

$x^*_{k+1} = x_k$

$\mathcal{X} = \{x \in \mathbb{R}^n \mid h(x) = 0, g(x) \leq 0\}$

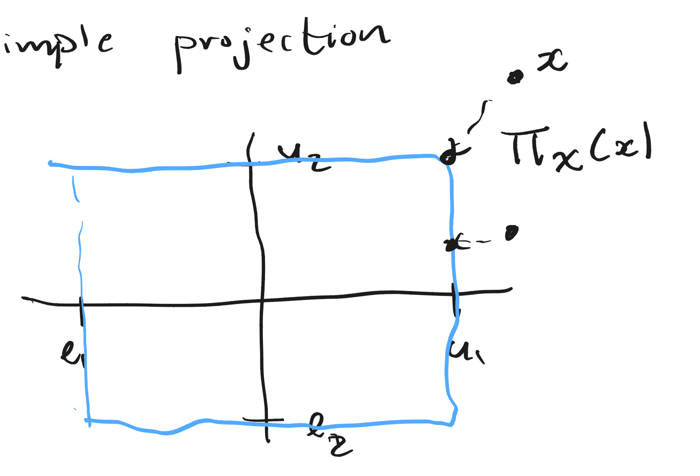
$\begin{bmatrix} h_1 \\ h_2 \\ \vdots \\ h_\ell \end{bmatrix}$

$$\Pi_{\mathcal{X}}(x_{unc}) = \arg\min_{x \in \mathcal{X}} \|x - x_{unc}\|_2$$

works well when $\mathcal{X}$ has a simple projection

<u>e.g.1</u> (box constraints)

$$\min f(x)$$
$$\text{s.t. } \ell \leq x \leq u$$



$$\begin{bmatrix} \ell_1 \\ \ell_2 \end{bmatrix} \leq \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$$

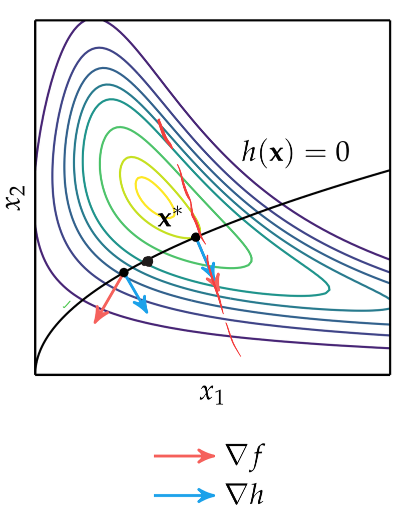
$$\Pi_{\mathcal{X}}(x)_i = \begin{cases} \ell_i, & x_i \leq \ell_i \\ x_i, & \ell_i \leq x_i \leq u_i \\ u_i, & \text{oth} \end{cases}$$

np.clamp(x, ℓ, u)

<u>Q</u>: How do we <u>define</u> constrained minima?

Consider a single equality constraint

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } h_1(x) = 0$$



An optimal point x^* requires:

$$\begin{cases} h_1(x^*) = 0 \\ \nabla f(x^*) = -\lambda \nabla h_1(x^*), \ \lambda \in \mathbb{R} \end{cases}$$

Lagrangian:

$$\mathcal{L}(x, \lambda) = f(x) + \lambda h(x)$$

$$\nabla_x \mathcal{L}(x^*, \lambda^*) = 0 = \nabla f(x^*) + \lambda^* \nabla h(x^*)$$
$$\nabla_\lambda \mathcal{L}(x^*, \lambda^*) = 0 = h_1(x^*)$$

For the multivariate case, $\lambda = [\lambda_1, \lambda_2, \dots, \lambda_\ell]$

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^T h(x)$$

<u>NOC:</u> $\nabla_x \mathcal{L}(x^*, \lambda^*) = 0$, $\nabla_\lambda \mathcal{L}(x^*, \lambda^*) = 0$

$$\nabla f(x^*) + \sum_{i=1}^{\ell} \lambda_i^* \nabla h_i(x^*) = 0$$

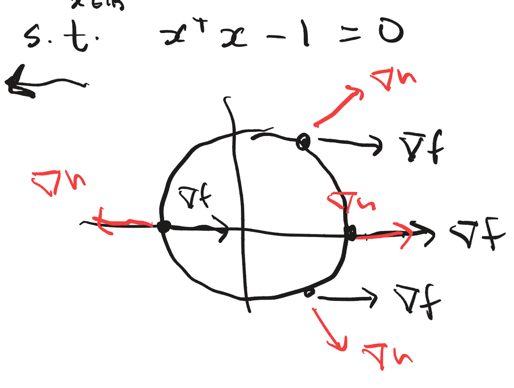
↑ linear combination

<u>e.g.1</u> (optim. on unit circle)

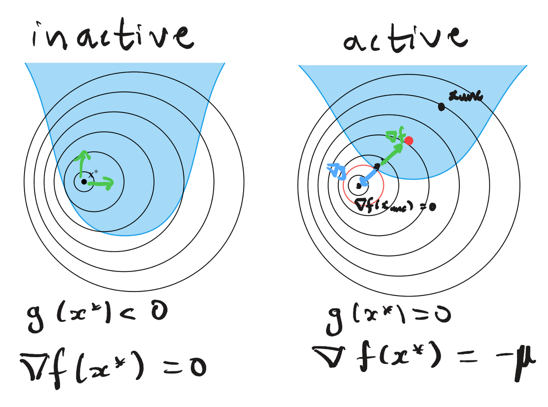
$$\min_{x \in \mathbb{R}^2} e_1^T x$$
$$\text{s.t. } x^T x - 1 = 0$$

$x^* = \begin{bmatrix} -1 \\ 0 \end{bmatrix}$ $\quad x^* \overset{?}{=} \begin{bmatrix} 1 \\ 0 \end{bmatrix}$

$\nabla h(x) = 2x$



Inequality constraints



inactive: $g(x^*) < 0$, $\nabla f(x^*) = 0$

active: $g(x^*) = 0$, $\nabla f(x^*) = -\mu \nabla g(x^*)$ for $\mu \geq 0$

$$\mathcal{L}(x, \mu) = f(x) + \mu g(x)$$

<u>NOC:</u>
$$\nabla_x \mathcal{L}(x^*, \mu^*) = 0$$
$$\nabla_\mu \mathcal{L}(x^*, \mu^*) = g(x^*) \leq 0$$
$$\mu \geq 0$$
$$\mu \cdot g(x) = 0$$

<u>Def</u> First-order constrained NOC (KKT)

For a constrained (smooth) problem

N&W Ch. 12

$$\min_{x \in \mathbb{R}^n} f(x)$$
$$\text{s.t. } h_i(x) = 0 \quad i = 1, \dots, \ell$$
$$g_j(x) \leq 0 \quad j = 1, \dots, m$$

For a minimum x^* of this problem, there must exist $\lambda^* \in \mathbb{R}^\ell$, $\mu^* \in \mathbb{R}^m$ satisfying

1. Stationarity
$$\nabla f(x^*) + \sum_{i=1}^{\ell} \lambda_i^* \nabla h_i(x^*) + \sum_{j=1}^{m} \mu_j \nabla g_j(x^*) = 0$$
2. Feasibility:
$$h(x^*) = 0$$
$$g(x^*) \leq 0$$
3. Dual feasibility: $\mu^* \geq 0$
4. Complementarity: $\mu^* \odot g(x^*) = 0$ $\quad (\mu_j^* = 0, \text{ or } g_j(x^*) = 0)$

These are sometimes called Karush-Kuhn-Tucker (KKT) conditions.

$$\mathcal{L}(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x)$$

$$\max_{\lambda, \mu \geq 0} \quad d(\lambda, \mu) = \arg\min_x \mathcal{L}(x, \lambda, \mu) \qquad \text{"dual"}$$

<u>Summary:</u>

when solving constrained problems:

1. Reformulate
2. Project $\Pi_{\mathcal{X}}$
3. Lagrangian / multipliers