2/5: Unconstrained optimization

Announcements:

- HW1 out now, due 2/12
- HW2 out 2/12

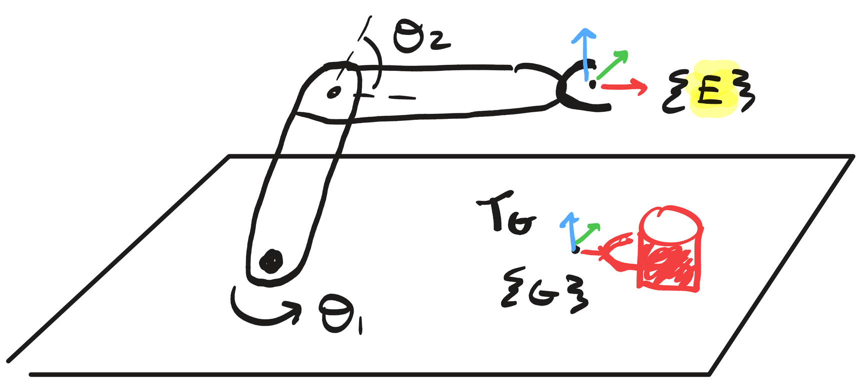
Last time:

- SO(3): $\oplus$, $\ominus$, Jacobians

Today:

- motivation: inverse kinematics (IK)
- unconstrained optimization
- descent methods

Problem: Inverse Kinematics (IK)



$q = [\theta_1, \theta_2, \dots, \theta_n] \in \mathbb{R}^n$

"configuration"

$x = \begin{bmatrix} R & p \\ \vec{0} & 1 \end{bmatrix} \in SE(3)$

"task variable"

"forward kinematics"

$FK(q): \mathbb{R}^n \rightarrow SE(3)$

$$\boxed{\min_{q \in \mathbb{R}^n} \| FK(q) \ominus T_G \|^2}$$

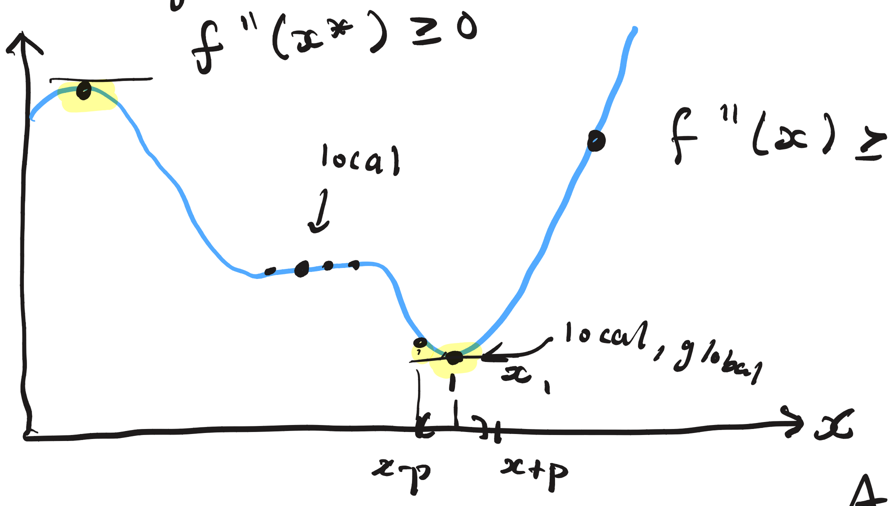
$$\min_{q \in \mathbb{R}^n} \left\| \begin{bmatrix} p_E^w - p_G^w \\ R_E \ominus R_G \end{bmatrix} \right\|^2$$

$$\boxed{\min_{q \in \mathbb{R}^n} f(q)}$$

Unconstrained minimization

$\min_{x \in \mathbb{R}^n} f(x)$, $f: \mathbb{R}^n \rightarrow \mathbb{R}$

"decision var" "objective function"



$f'(x) = 0$

$f''(x^*) \geq 0$

$f''(x) \geq 0$

Def 1 (Local minimum)

A point $x^* \in \mathbb{R}^n$ is a local min. of f if $\exists\, \varepsilon > 0$ s.t.

$f(x^* + p) \geq f(x^*)$ $p.\ \|p\| < \varepsilon$

Def 1 (global minimum)

A point x^* is a global min if $f(x) \geq f(x^*)$ $\forall x \in \mathbb{R}^n$

Prop 1 (First-order NOC)

A local min x^* of f must satisfy $\nabla f(x^*) = 0$

Prop 1 (second-order NOC)

A local min x^* of f must satisfy $\nabla^2 f(x^*) \succeq 0$.

$A \succeq 0$

$\forall x \in \mathbb{R}^n$

$x^T A x \geq 0$

Descent methods

Alg 1 (Descent method)

input: initial guess x_0 ①

for $k = 1, 2, \dots, K_{max}$

  $d_k \leftarrow$ descent_dir(x_k) ?

  $\alpha_k \leftarrow$ step_size(x_k, d_k) ✓

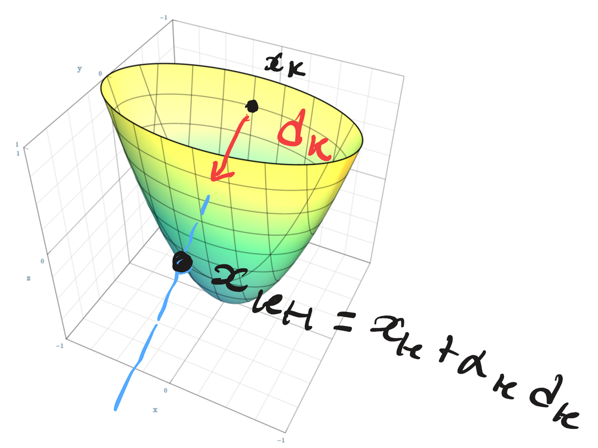
  $x_{k+1} \leftarrow x_k + \alpha_k d_k$

  if terminate(x_{k+1}): ✓

    break

$\|\nabla f(x_{k+1})\| < \varepsilon$ ②

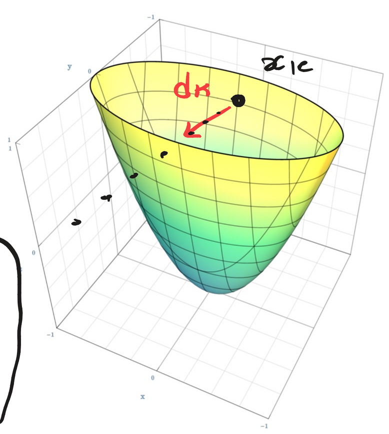
$f(x_k) - f(x_{k+1}) < \varepsilon$ ③

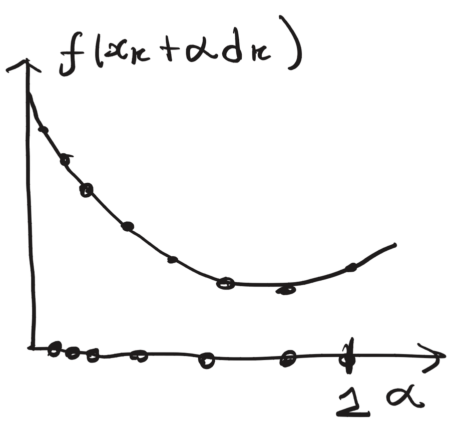


Line search

"parallel line search"

$$\boxed{\vec{\alpha} = \text{logspace}(\underset{1e-3}{-3}, \underset{1e0}{0}, \underset{\#}{15})}$$

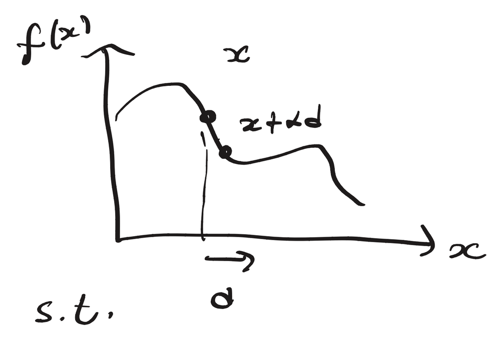




$\text{vmap}(f, x, d)(\vec{\alpha}) = f^{(i)} \rightarrow$ choose $\arg\min_i f(x + \alpha^i d)$

$O(1)$

Descent directions



Def 1 (Descent direction)

A vector $d \in \mathbb{R}^n$ is a descent direction for f at x if $\exists\, \alpha > 0$ s.t.

$f(x + \alpha d) < f(x)$

Fact: d is a descent direction if $d^T \nabla f(x) < 0$

Idea 1: Gradient descent $\boxed{d = -\nabla f(x)}$

$d^T \nabla f(x) = -\nabla f(x)^T \nabla f(x) < 0$

Idea 2: Newton's method

$f(x+d) \approx f(x) + \nabla f(x)^T d + \frac{1}{2} d^T \nabla^2 f(x) d + \cancel{O(\|d\|^3)}^{\,0}$

$\min_d\ q(d) = f(x) + g^T d + \frac{1}{2} d^T H d$, $g = \nabla f(x)$

$H = \nabla^2 f(x)$

if $H \succ 0$, d^* would satisfy

$\nabla_d q(d^*) = 0 \Rightarrow g + Hd^* = 0$

$\Rightarrow \boxed{Hd = -g}$ $\leftarrow$ trading off convergence speed for numerical stability.