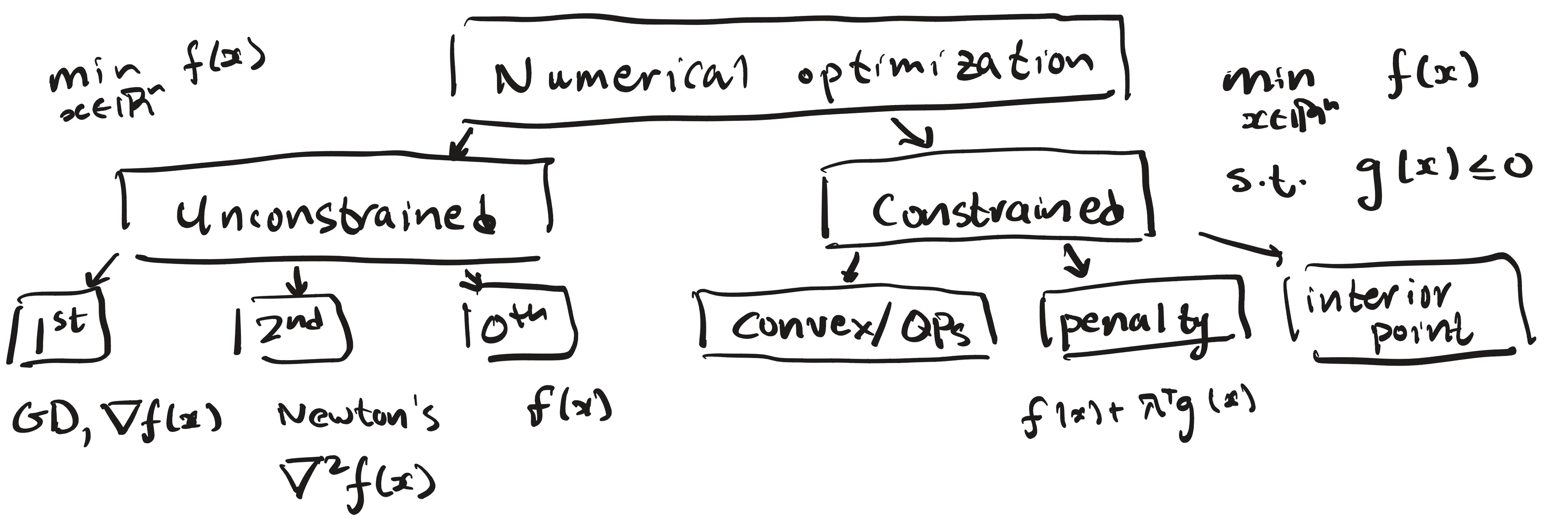
2/12: Constrained Optimization

Announcements

- HW1 due today, HW2 out ASAP.
- Project proposal due next Thurs
  - 1 pg. max, project idea(s)
  - groups of ≤ 3
  - apply tools from class to fun problem



Last time: Nonlinear LS

$$f(x) = \frac{1}{2}\sum_{i=1}^{N} r_i(x)^2 \qquad r_i(x): \mathbb{R}^n \to \mathbb{R}$$

$$\nabla_x f(x) = J(x)^T r(x) \qquad r(x) = [r_1(x), r_2(x), \dots, r_N(x)]$$

$$\nabla_x^2 f(x) = J(x)^T J(x) + \underbrace{\left[\sum_{i=1}^{N} r_i(x)\nabla^2 r_i(x)\right]}_{\approx 0} = 0$$

$$\boxed{J^T J d = -J^T r}$$ Gauss-Newton "normal equations"

$$d \overset{?}{=} -(J^T J)^{-1} J^T r$$

Pros: – only requires 1st order ($J(x)$)
– can leverage sparsity of J

Cons: $O(n^3)$ compute
$O(n^2)$ memory

Idea! Matrix-free solver

$\underset{n\times n}{Ax = b} \Rightarrow$ matvec(v): Av

np.linalg.solve(A, b) — matrix $\quad$ scipy.sparse.linalg.cg(matvec, b) $\quad O(n)$

jax.jacobian

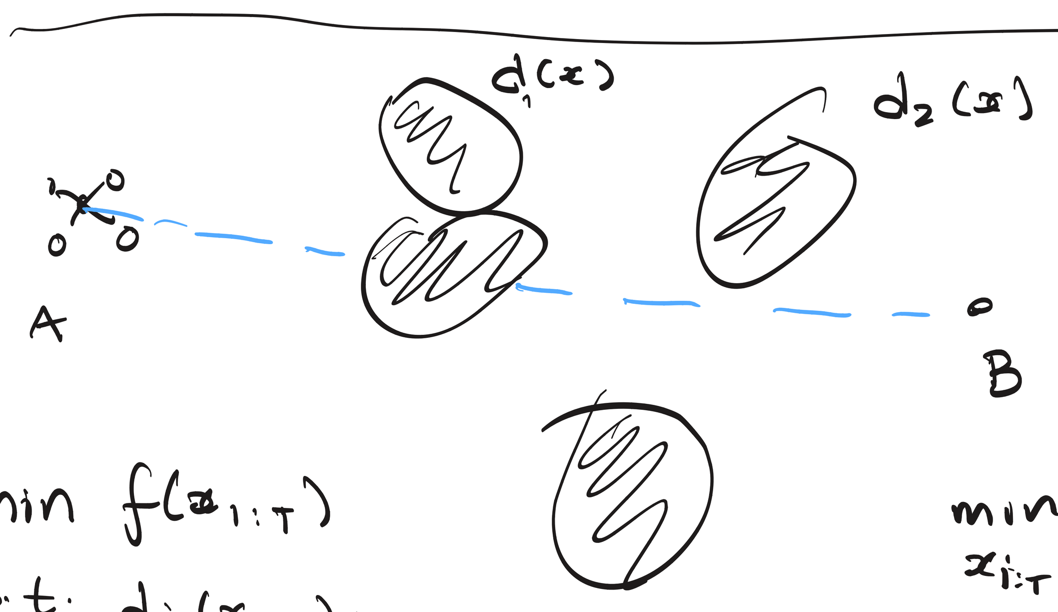
Autodiff: $r: \mathbb{R}^n \to \mathbb{R}^m \quad J = \frac{\partial r}{\partial x} \in \mathbb{R}^{m\times n} = [\text{jvp}(e_1) \dots \text{jvp}(e_n)]$

JVP: $v \mapsto Jv$ (forward) $\qquad$ VJP: $u \mapsto J^T u$ (reverse)

$$J^T J d = -J^T r \Rightarrow J^T(Jd) = -J^T r$$

$$\underbrace{\text{vjp}(\text{jvp}(d))}_{} = -\text{vjp}(r)$$

$$\frac{\partial r}{\partial x} = \overset{u^T}{(J_N \cdot J_{N-1} \cdot \dots \cdot J_1)} v$$



$$\min f(x_{1:T}) \quad \text{s.t. } \underbrace{d_i(x_{1:T}) \ge 0}_{} \quad \forall t, i$$

$$\min_{x_{1:T}} f(x_{1:T}) + \rho \sum_{i,t} d_i(x_t)$$

$\rho \to \infty$

One strategy: reparameterize

e.g. $SO(3)$

$$\min_{R \in SO(3)} f(R) \Rightarrow \min_{R \in \mathbb{R}^{3\times 3}} f(R) \text{ s.t. } R^T R = I,\ \det(R) = +1 \Rightarrow \min_{\theta \in \mathbb{R}^3} \underbrace{f(\exp(\theta))}_{}$$

$$\min_{x \in \mathbb{R}} f(x) \text{ s.t. } x \ge 0 \Rightarrow \min_{t \in \mathbb{R}} f(e^t)$$

$$\min_{M \in \mathbb{R}^{n\times n}} f(M) \text{ s.t. } M \succeq 0 \Rightarrow \min_{L \in \mathbb{R}^{n\times n}} f(L^T L)$$

TPS: $\displaystyle \min_{x \in \mathbb{R}^n} f(x) \quad \text{s.t. } \sum_{i=1}^{n} a_i x_i = b$

hint: $f(x) = f(x_1, x_2, \dots, x_n)$

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

$$x_n = \left(b - \sum_{i=1}^{n-1} a_i x_i\right)/a_n = \hat{x}_n \Rightarrow \min_{\tilde{x} \in \mathbb{R}^{n-1}} \underbrace{f\left(x_1, \dots, \frac{b - \sum a_i x_i}{a_n}\right)}_{}$$