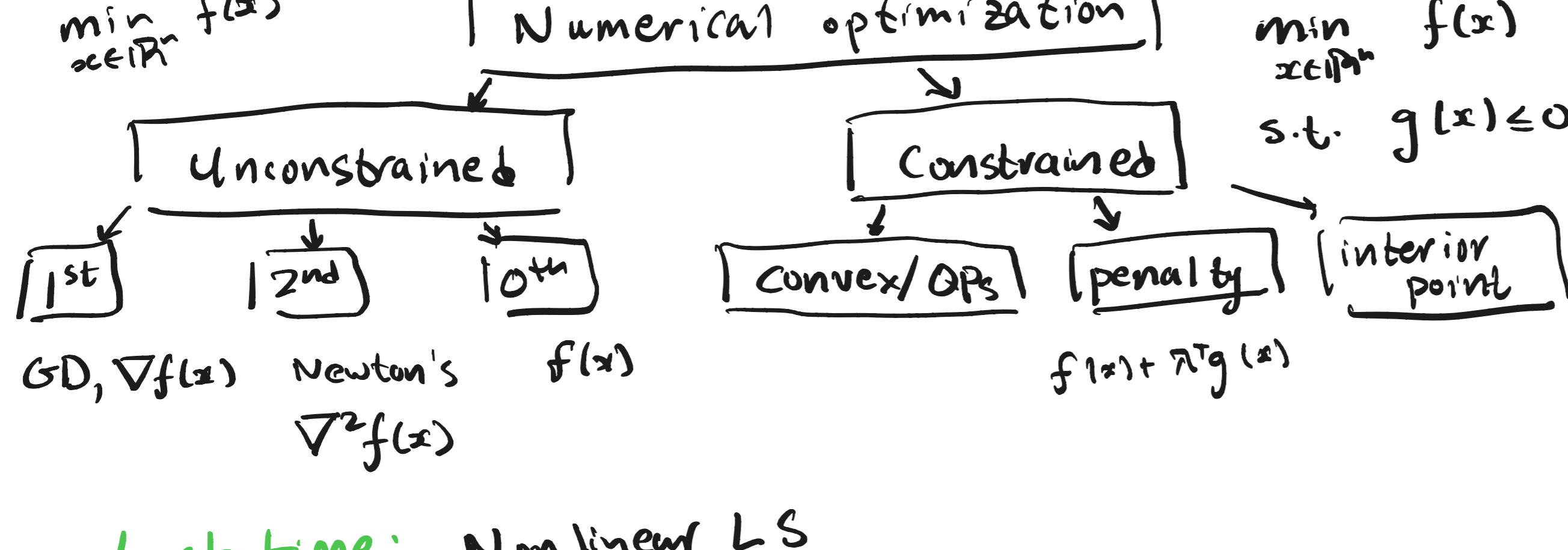


2/12: Constrained Optimization

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

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

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

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

$$J^T J d = -J^T r \quad \text{Gauss-Newton "normal equations"} \\ d = -(J^T J)^{-1} J^T r$$

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

$$Ax = b \Rightarrow \text{matvec}(v) : Av$$

$n \times n$

`np.linalg.solve(A, b)` `scipy.sparse.linalg.cg(matvec, b)`
 matrix jac, jacobian $O(n)$

$$\text{Autodiff: } r : \mathbb{R}^n \rightarrow \mathbb{R}^m \quad J = \frac{\partial r}{\partial x} \in \mathbb{R}^{m \times n} = [j_{ij}(x) \dots j_{jn}(x)]$$

$$\begin{array}{ll} \text{JVP: } v \mapsto Jv & \text{VJP: } u \mapsto J^T u \\ \text{forward} & \text{reverse} \end{array}$$

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

$$v \text{JVP}(j \text{JVP}(d)) = -v \text{JVP}(r)$$

$$\frac{\partial r}{\partial x} = \underbrace{(J_N \cdot J_{N-1} \cdots J_1)}_{\sim \sim} v$$

A sequence of nested ellipses representing a trajectory. The first ellipse is labeled x_0 . Subsequent ellipses are labeled $d_1(x)$, $d_2(x)$, and so on. Arrows indicate the progression from x_0 through the ellipses to a final point B .

$$\min f(x_{1:T})$$

$$\text{s.t. } d_i(x_{1:T}) \geq 0 \quad \forall t, i$$

$$\min f(x_{1:T}) + p \sum_{i=1}^T d_i(x_t)$$

$$p \rightarrow \infty$$

One strategy: reparameterize

e.g. $\text{SO}(3)$

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

$$\min_{x \in \mathbb{R}^n} f(x) \Rightarrow \min_{t \in \mathbb{R}} f(e^t)$$

s.t. $x \geq 0$

$$\min_{M \in \mathbb{R}^{n \times n}} f(M) \Rightarrow \min_{L \in \mathbb{R}^{n \times n}} f(L^T L)$$

s.t. $M \succeq 0$

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

s.t. $\sum_{i=1}^n q_i x_i = b$

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

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

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