IMPORTANT ANNOUNCEMENT:

- Final Exam on Wednesday, Dec 4, 2024 (last ESE 415 class)
- Two lectures preceding final exam: review & exam preparation

Convex Optimization Problems

A convex optimization problem is one of the form

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq b_i \quad , \; i=1,\dots,m \end{array} \qquad \text{(COP)}$$

where the functions $f_0, f_1, \dots, f_m : \mathbb{R}^n \to \mathbb{R}$ are convex.

Comments:

- Least squares & linear programming problems are special cases of the general convex optimization problem (COP)
- There is in general no analytical formula for the sol'n of (COP), but there are very efficient computational methods for solving them, e.g., interior-point methods (can easily solve (COP) w/ 100s of variables and 1000s of constraints on a desktop computer, in at most a few seconds.)
- Using convex optimization is very much like using least-squares & linear programs. If we can formulate a problem as a (COP), we have basically solved the original problem.

Practical Methods for establishing that a function is convex:

1. very definition $f(\alpha x+(1-\alpha)y) \leq \alpha f(x)+(1-\alpha)f(y)$, often simplified by restricting to a line
2. for twice differentiable functions: show $(Hf)(x) \succeq 0$.
3. Show that f is obtained from simple convex functions by operations that preserve convexity:
   - nonnegative weighted sums of convex fcn's are convex (TBD)
   - composition of affine functions
   - pointwise maximum and supremum
   - certain other compositions (HW4, P5)

there exist toolboxes for checkings automatically

Theorem: If $f_1, f_2 \in \mathcal{F}^1(C)$ and $a,b \geq 0$, then $f = af_1 + bf_2 \in \mathcal{F}^1(C)$.

Proof: Since $f_1, f_2 \in \mathcal{F}_1$: $\forall x,y \in C$

$$f_1(y) \geq f_1(x) + (\nabla f_1)(x)^T(y-x)$$
$$f_2(y) \geq f_2(x) + (\nabla f_2)(x)^T(y-x)$$

$$\Rightarrow f(y) = a f_1(y) + b f_2(y)$$
$$\geq a\left(f_1(x) + (\nabla f_1)(x)^T(y-x)\right) + b\left(f_2(x) + (\nabla f_2)(x)^T(y-x)\right)$$
$$= f(x) + \underbrace{\left[a(\nabla f_1)(x)^T + b(\nabla f_2)(x)^T\right](y-x)}_{=(\nabla f)(x)^T(y-x)}$$

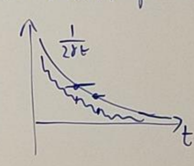
$$\nabla(\alpha f_1 + b f_2) = \alpha(\nabla f_1) + b(\nabla f_2)$$

$$\Rightarrow f \in \mathcal{F}^1(C) \quad \blacksquare$$

Next time: Gradient Method (GM) for (unconstrained) convex optimization problems

Theorem: Let $f \in \mathcal{F}^1_L(\mathbb{R}^n)$ with finite minimum $f^* = f(x^*)$. Then, for any step size $0 \leq \gamma \leq \frac{1}{L}$, the iterates of the GM satisfy

$$f(x^t) - f(x^*) \leq \left(\frac{1}{2\gamma t}\right)\underbrace{\|x^0 - x^*\|^2}_{\text{const.}}.$$



$O(1/t)$

Proof: Next lecture

Q: What is the convergence rate of the sequence $\{f(x^t)\}_{t=0,1,2,\ldots}$?