Outline

Mathematical Background

Lecture 4: Intro to Optimization
Lecture 5: Gradient Descent

Lecture 4: Intro to Optimization

Unconstrained Optimization
- First Order Optimality
- Second Order Optimality

Convexity
- What is convexity?
- Convex optimization

Constrained Optimization
- Lagrangian
- Examples
Most Optimization Problems are Not Easy

Minimize the log-sum-exp function:

\[ f(x) = \log \left( \sum_{i=1}^{m} \exp(a_i^T x + b_i) \right) \]

- Gradient is (exercise)

\[ \nabla f(x^*) = \frac{1}{\sum_{j=1}^{m} \exp(a_j^T x^* + b_j)} \sum_{i=1}^{m} \exp(a_i^T x^* + b_i) a_i. \]

- Non-linear equation. No closed-form solution.
- Need iterative algorithms, e.g., gradient descent.
- Or off-the-shelf optimization solver, e.g., CVX.
Disciplined optimization: It translates the problem for you.
Developed by S. Boyd and colleagues (Stanford).
E.g., Minimize $f(x) = \log \left( \sum_{i=1}^{n} \exp(a_i^T x + b_i) \right) + \lambda \| x \|^2$. 

```python
import cvxpy as cp
import numpy as np

n = 100
d = 3
A = np.random.randn(n, d)
b = np.random.randn(n)
lambda_ = 0.1

x = cp.Variable(d)
objective = cp.Minimize(cp.log_sum_exp(A*x - b) + lambda_*cp.sum_squares(x))
constraints = []
prob = cp.Problem(objective, constraints)

optimal_objective_value = prob.solve()
print(optimal_objective_value)
print(x.value)
```
Convex Function

Definition

Let $\mathbf{x} \in \mathcal{X}$ and $\mathbf{y} \in \mathcal{X}$. Let $0 \leq \lambda \leq 1$. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is convex over $\mathcal{X}$ if

$$f(\lambda \mathbf{x} + (1 - \lambda) \mathbf{y}) \leq \lambda f(\mathbf{x}) + (1 - \lambda) f(\mathbf{y}).$$

The function is called strictly convex if “≤” is replaced by “<”.
Example: Which one is convex?
Verifying Convexity

Any of the following conditions is necessary and sufficient for convexity:

1. By definition:

   \[ f(\lambda x + (1 - \lambda)y) \leq \lambda f(x) + (1 - \lambda)f(y). \]

   - Function value is lower than the line.

2. First Order Convexity:

   \[ f(y) \geq f(x) + \nabla f(x)^T (y - x), \quad \forall x, y \in \mathcal{X}. \]

   - Tangent line is always lower than the function

3. Second Order Convexity: \( f \) is convex over \( \mathcal{X} \) if and only if

   \[ \nabla^2 f(x) \succeq 0 \quad \forall x \in \mathcal{X}. \]

   - Curvature is positive.
Tangent Line Condition Illustrated

\[ f(y) = f(x) + \nabla f(x)^T (y-x) \]