Today’s Topics

• Convexity and Convex Optimization
• Duality and Optimality Conditions
• Review
Establish Convexity

Assume $\text{dom} f$ is convex.

Restricting to line: $f(\cdot)$ is convex $\Leftrightarrow$ for all $x \in \text{dom} f$ and for all $v$

$g(t) = f(x + tv)$ is convex on its domain $\{t \in \mathbb{R} : x + tv \in \text{dom} f\}$

Differentiable $f(\cdot)$ with a convex domain is convex iff

$f(y) \geq f(x) + \nabla f(x)^T(y - x)$ for all $x, y \in \text{dom} f$

$\Rightarrow$ first order approximation is a global underestimator of $f(\cdot)$


**Proof of First-order condition**

Proof: Consider the case n = 1:

We prove that a differentiable function $f : \mathbb{R} \to \mathbb{R}$ is convex iff

$$f(y) \geq f(x) + f'(x)(y - x)$$

(1)

for all $x, y \in \text{dom} f$.

⇒ Assume that $f$ is convex and $x, y \in \text{dom} f$. Since $\text{dom} f$ is convex (i.e., an interval), for all $0 < t < 1$, $x + t(y - x) \in \text{dom} f$, and by convexity of $f$,

$$f(x + t(y - x)) \leq (1 - t)f(x) + tf(y).$$

If we divide both sides by $t$, we obtain

$$f(y) \geq f(x) + \frac{f(x + t(y - x)) - f(x)}{t}.$$
and taking the limit as $t \to 0$ yields (1).

$L$ Assume the function satisfies (1) for all $x, y \in \text{dom} f$ (which is an interval).

Choose any $x \neq y$, and $0 \leq \theta \leq 1$, and let $z = \theta x + (1 - \theta)y$. Applying (1) twice yields

$$f(x) \geq f(z) + f'(z)(x - z), \quad f(y) \geq f(z) + f'(z)(y - z).$$

Multiplying the first inequality by $\theta$, the second by $1 - \theta$, and adding them yields

$$\theta f(x) + (1 - \theta)f(y) \geq f(z), \quad i.e., \quad f \text{ is convex}$$
1-D Examples

convex:

- $ax + b$ on $R$, for any $a, b \in R$
- $e^{ax}$, for any $a \in R$
- $x^\alpha$ when $x > 0$, for $\alpha \geq 1$
- $|x|^p$ on $R$, for $p \geq 1$

concave:

- $ax + b$ on $R$, for any $a, b \in R$
- $x^\alpha$ when $x > 0$, for $0 \leq \alpha \leq 1$
- $\log(x)$ when $x > 0$
Examples on $\mathbb{R}^n$

- affine function $f(x) = a^T x + b$
- norms: $\|x\|_p = (\sum_{i=1}^{n} |x_i|^p)^{1/p}$ for $p \geq 1$; $\|x\|_\infty = \max_k |x_k|$
Positively weighted sum and composition

- $\sum_{i=1}^{m} \alpha_i f_i(x)$ is convex if $\{f_i(x)\}$ are convex and $\alpha_i \geq 0$ (extends to infinite sums, integrals)

- $f(Ax + b)$ is convex if $f(\cdot)$ is convex.

Example: norm of affine function: $f(x) = \|Ax + b\|$
Maximum and Supremum

• If \( f_1(x), \ldots, f_m(x) \) are convex, then \( f(x) = \max\{f_1(x), \ldots, f_m(x)\} \) is convex.

  Example: \( f(x) = \max_i (a_i^T x - b_i) \)

• If \( f(x, y) \) is convex in \( x \) for each \( y \in \mathcal{D}_y \), then \( g(x) = \sup_{y \in \mathcal{D}_y} f(x, y) \) is convex.

  Example: \( f(x) = \max_{y \in \mathcal{D}_y} \|x - y\| \)
Proof:

Let $f(x) = \max(f_1(x), f_2(x))$.

Note $\max(a + b, c + d) \leq \max(a, c) + \max(b, d)$

Assume that $f_1(\cdot)$ and $f_2(\cdot)$ are convex. Then

$$f_1(\theta x + (1 - \theta)y) \leq \theta f_1(x) + (1 - \theta)f_1(y)$$
$$f_2(\theta x + (1 - \theta)y) \leq \theta f_2(x) + (1 - \theta)f_2(y)$$

Then

$$f(\theta x + (1 - \theta)y) = \max(f_1(\theta x + (1 - \theta)y), f_2(\theta x + (1 - \theta)y)$$
$$\leq \max(\theta f_1(x) + (1 - \theta)f_1(y), \theta f_2(x) + (1 - \theta)f_2(y))$$
$$\leq \max(\theta f_1(x), \theta f_2(x)) + \max((1 - \theta)f_2(x), (1 - \theta)f_2(y))$$
$$= \theta f(x) + (1 - \theta)f(y)$$
Lagrangian and Optimality

• **Lagrangian:** \( L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^l \to \mathbb{R}, \)

\[
L(x, \lambda, \nu) \overset{\text{def}}{=} f(x) + \sum_{i=1}^{m} \lambda_i g_i(x) + \sum_{j=1}^{l} \nu_j q_j(x)
\]

• \( \lambda_i \): Lagrange multiplier associated with \( g_i(x) \leq 0 \)

• \( \nu_j \): Lagrange multiplier associated with \( q_j(x) = 0 \)
Karush-Kuhn-Tucker (KKT) Conditions

Assume $f(\cdot), q(\cdot), g(\cdot)$ are \textbf{differentiable}.

$$\min_x f(x)$$

subject to \quad $q_i(x) = 0, \quad i = 1, \cdots, l$ \quad \text{(NLP)}

$$g_i(x) \leq 0, \quad i = 1, \cdots, m$$

Assume strict duality holds at feasible $(x^*, \lambda^*, \nu^*)$

$$f(x^*) = G(\lambda^*, \nu^*) = \inf_x L(x, \lambda^*, \nu^*)$$

$$\leq f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{j=1}^{l} \nu_j^* q_j(x^*)$$

$$\leq f(x^*) \quad \text{(due to feasibility)}$$

$\Rightarrow$ both inequalities must be equalities
The first inequality becomes equality,

\[
\left( \inf_x L(x, \lambda^*, \nu^*) \right) = f(x^*) + \sum_{i=1}^{m} \lambda_i^* g_i(x^*) + \sum_{j=1}^{l} \nu_j^* q_j(x^*)
\]

Hence \( x^* \) minimizes \( L(x, \lambda^*, \nu^*) \). This implies

\[
\nabla_x L(x^*, \lambda^*, \nu^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{l} \nu_j^* \nabla q_j(x^*) = 0
\]

The second inequality becomes equality implies complementarity

conditions \( \lambda_i^* g_i(x^*) = 0 \) for all \( i \)
Karush-Kuhn-Tucker (KKT) Optimality

KKT Conditions: Let $f, q, g$ be continuously differentiable.

- **Primal Feasibility:**
  \[ q_i(x^*) = 0, \quad i = 1, \cdots, l, \quad g_i(x^*) \leq 0, \quad i = 1, \cdots, m \]

- **Dual Feasibility:** $\lambda^* \geq 0$

- **Gradient with respect to $x$, $\nabla_x L(\cdot)$, equal zero:**

  \[
  \nabla_x L(x^*, \lambda^*, \nu^*) = \nabla f(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla g_i(x^*) + \sum_{j=1}^{l} \nu_j^* \nabla q_j(x^*) = 0
  \]

  \hspace{1cm} (2)

- **Complementarity:** $\lambda_i^* g_i(x^*) = 0, \quad i = 1, \cdots, m$