Homework #2

Due Mar 1 (Wednesday) at the beginning of class
Please show all work and intermediate steps. Late submission will lead to 0 credit.

Problem 1: Subgradient and Subdifferential

Exercise 1.1 (Subdifferential) Calculate $\partial f(x)$ for the following functions

(a) $f(x) = \max(1, |x| - 1)$ on $\mathbf{R}$.

(b) $f(x) = \|x\|$, where $\|\cdot\|$ is a norm on $\mathbf{R}^n$.

Solution

(a)

$$\partial f(x) = \begin{cases} \text{sgn}(x) & |x| > 2 \\ [-1, 0] & x = -2 \\ [0, 1] & x = 2 \\ \{0\} & |x| < 2 \end{cases}$$

(b)

$$\partial f(x) = \{g : g^T x = \|x\| \text{ and } \|g\|_* \leq 1\}$$

Exercise 1.2 (Subdifferential of Pointwise Maximum) Let $f_1, \ldots, f_m$ be convex functions and $x \in \cap_{i=1}^m \text{int}(\text{dom}(f_i))$. Let $h(x) = \max_{1 \leq i \leq m} f_i(x)$, and $I(x)$ be the set of all $i \in \{1, \ldots, m\}$ such that $f_i(x) = h(x)$.

(a) Use definition of subdifferential to show that $\text{Conv}(\{\partial f_i(x) : i \in I(x)\}) \subseteq \partial h(x)$.

(b) Use definiton of directional derivative to show that $h'(x; d) = \max_{i \in I(x)} f_i'(x; d)$.

(c) Use separation theorem to show that $\text{Conv}(\{\partial f_i(x) : i \in I(x)\}) \supseteq \partial h(x)$.
[*Hint: Use the fact that $\partial f_i(x)$ are compact convex and $h'(x; d) = \max_{g \in \partial h(x)} g^T d$.*]

Solution

(a) By definition of subdifferential, we have

$$\forall i \in I(x), g \in \partial f_i(x) : f_i(y) \geq f_i(x) + g^T(y - x), \forall x, y$$

By definition of $I(x)$, we know that $f_i(x) = h(x), \forall i \in I(x)$. Hence

$$\forall i \in I(x), g \in \partial f_i(x) : h(y) \geq h(x) + g^T(y - x), \forall x, y, \quad \text{i.e. } g \in \partial h(x)$$

Hence, $\{\partial f_i(x) : i \in I(x)\} \subset \partial h(x)$. Since $\partial h(x)$ is a convex set, we further have $\text{Conv}(\{\partial f_i(x) : i \in I(x)\}) \subseteq \partial h(x)$.

(b) By definition of directional derivative, we have

$$\begin{aligned} h'(x;d) &= \lim_{t\to 0_+} \frac{h(x+td)-h(x)}{t} \\ &= \lim_{t\to 0_+} \frac{\max_{i\in I(x)} f_i(x+td)-h(x)}{t} \\ &= \max_{i\in I(x)} \lim_{t\to 0_+} \frac{f_i(x+td)-h(x)}{t} \\ &= \max_{i\in I(x)} \lim_{t\to 0_+} \frac{f_i(x+td)-f_i(x)}{t} \\ &= \max_{i\in I(x)} f_i'(x;d) \end{aligned}$$

(c) Let $C = \text{Conv}(\{\partial f_i(x) : i \in I(x)\})$. Since $\partial f_i(x)$ is compact for any $i \in I(x)$ and $x \in \cap_{i=1}^m \text{int}(\text{dom}(f_i))$, so C must be compact convex. Suppose $C \subsetneq \partial h(x)$, then there exists $g \in \partial h(x)$ and $g \neq C$. By Separation Theorem, there exists a vector $d \neq 0$, such that $d^T g > \sup_{g\in C} d^T g$. This further implies that

$$d^T g > \sup_{g\in C} d^T g \geq \max_{i\in I(x)} \max_{g\in\partial f_i(x)} d^T g = \max_{i\in I(x)} f_i'(x;d) = h'(x;d) = \max_{g\in\partial h(x)} d^T g$$

which leads to a contradiction. Therefore, $\text{Conv}(\{\partial f_i(x) : i \in I(x)\}) \supseteq \partial h(x)$.

Combining with (a) and (c), we have that

$$\partial h(x) = \text{Conv}(\{\partial f_i(x) : i \in I(x)\}).$$

Exercise 1.3 (Directional Derivative) Let f be a convex function and $x \in \text{dom}(f)$ and let d be such that $x + \alpha d \in \text{dom}(f)$ for $\alpha \in (0, \delta)$ for some $\delta > 0$. Show that the scalar function

$$\phi(\alpha) = \frac{f(x+\alpha d) - f(x)}{\alpha}$$

is non-decreasing function of α on $(0, \delta)$.

Solution Let $\alpha_2 > \alpha_1 > 0$. Note that $x + \alpha_1 d = \frac{\alpha_1}{\alpha_2}(x + \alpha_2 d) + (1 - \frac{\alpha_1}{\alpha_2})x$ By convexity of f, we have

$$f(x+\alpha_1 d) = \frac{\alpha_1}{\alpha_2} f(x+\alpha_2 d) + (1 - \frac{\alpha_1}{\alpha_2}) f(x).$$

Rearranging terms leads to

$$\frac{f(x+\alpha_1 d) - f(x)}{\alpha_1} \leq \frac{f(x+\alpha_2 d)}{\alpha_2}.$$

Problem 2: Convex Conjugate

Exercise 2.1 (Compute Conjugate) Calculate the conjugate of the following functions:

(a) $f(x) = e^x$ on $\mathbf{R}$

(b) $f(x) = \|x\|$ on $\mathbf{R}^n$

(c) $f(x) = \frac{1}{2}\|x\|^2$ on $\mathbf{R}^n$

(d) $f(x) = \log(\sum_{i=1}^n \exp\{x_i\})$ on $\mathbf{R}^n$

Solution

(a) The conjugate for $f(x) = e^x$ is

$$f^*(y) = \begin{cases} y \ln y - y, & y > 0 \\ 0, & y = 0 \\ +\infty, & \text{o.w.} \end{cases}$$

(b) The conjugate for $f(x) = \|x\|$ is

$$f^*(y) = \begin{cases} 0, & \|y\|_* \leq 1 \\ +\infty, & \text{o.w.} \end{cases}$$

(c) The conjugate for $f(x) = \frac{1}{2}\|x\|^2$ is

$$f^*(y) = \frac{1}{2}\|y\|_*$$

(d) The conjugate for $f(x) = \log(\sum_{i=1}^n \exp\{x_i\})$ is

$$f^*(y) = \begin{cases} \sum_{i=1}^n y_i \log(y_i), & \text{if } y \geq 0 \text{ and } \sum_{i=1}^n y_i = 1 \\ +\infty, & \text{o.w.} \end{cases}$$

Exercise 2.2 (Calculus of Conjugate) Prove the following

(a) (Scalar Multiplication) Let $f(x)$ be convex and $\alpha > 0$, then

$$(\alpha f)^*(y) = \alpha f^*(y/\alpha)$$

(b) (Direct Summation) Let $f(x_1)$ and $g(x_2)$ be convex and $h(x_1, x_2) = f(x_1) + g(x_2)$, then

$$h^*(y_1, y_2) = f^*(y_1) + g^*(y_2)$$

(c) (Weighted Summation) Let $f(x)$ and $g(x)$ be closed convex functions,and $h(x) = f(x) + g(x)$, then

$$h^*(y) = \inf_z \{f^*(z) + g^*(y - z)\}$$

where the latter is the convolution of f^* and g^*.
[*Hint: First show that* $(\inf_z\{F(z) + G(y - z)\})^* = F^*(y) + G^*(y)$, and then apply with $F = f^*$, and $G = g^*$.]

Solution

(a) Scalar Multiplication:

$$(\alpha f)^*(y) = \sup_x \{y^T x - \alpha f(x)\} = \alpha \sup_x \{(y/\alpha)^T x - f(x)\} = \alpha f^*(y/\alpha)$$

(b) Direct Summation:

$$\begin{aligned} h^*(y_1, y_2) &= \sup_{x_1, x_2} \{y_1^T x_1 + y_2^T x_2 - f(x_1) - g(x_2)\} \\ &= \sup_{x_1}\{y_1 T x_1 - f(x_1)\} + \sup_{x_2}\{y_2 T x_2 - g(x_2)\} \\ &= f^*(y_1) + g^*(y_2) \end{aligned}$$

(c) Weighted Summation: first, we prove the following equality $(F\Box G)^*(x) = F^*(x) + G^*(x)$ where $F\Box G$ denotes the convlution operator $F\Box G = \inf_y \{F(y) + G(x-y)\}$. This is because

$$\begin{aligned}
(F\Box G)^*(x) &= \sup_z \left\{ z^T x - \inf_y (F(y) + G(z-y)) \right\} \\
&= \sup_z \left\{ z^T x - \inf_{y_1+y_2=z} (F(y_1) + G(y_2)) \right\} \\
&= \sup_z \left\{ \sup_{y_1+y_2=z} \{(y_1+y_2)^T x - F(y_1) - G(y_2)\} \right\} \\
&= \sup_{y_1,y_2} \{(y_1+y_2)^T x - F(y_1) - G(y_2)\} \\
&= \sup_{y_1} \{y_1^T x - F(y_1)\} + \sup_{y_2} \{{y_2}^T x - G(y_2)\} \\
&= F^*(x) + G^*(x)
\end{aligned}$$

Using $F = f^*$, and $G = g^*$, and the fact that $F^* = f$ and $G^* = g$, this leads to

$$(f^*\Box g^*)^*(x) = f(x) + g(x)$$

Note that $f + g$ is closed, so taking conjugate on both sides will still hold, i.e. $f^*\Box g^*$ is closed, hence,

$$(f^*\Box g^*)^{**}(x) = (f(x) + g(x))^*$$

We can also easily show that the convolution of two closed functions is still closed, i.e. $f^*\Box g^* = (f^*\Box g^*)^{**}$. Combining these these two facts, we arrive at

$$(f+g)^*(y) = (f^*\Box g^*)(y) = \inf_z \{f^*(z) + g^*(y-z)\}$$

Exercise 2.3 (Fenchel's Inequality) We already know that for any x and y, $x^T y \leq f(x) + f^*(y)$. Show that $x^T y = f(x) + f^*(y)$ if and only if $y \in \partial f(x)$.

Solution This is because

$$\begin{aligned}
x^T y = f(x) + f^*(y) &\iff x^T y - f(x) = \sup_z \{z^T y - f(z)\} \\
&\iff x^T y - f(x) \geq z^T y - f(z), \forall z \\
&\iff f(z) \geq f(x) + y^T (z-x), \forall z \\
&\iff y \in \partial f(x)
\end{aligned}$$

Problem 3: Lagrange Duality

Consider the following optimization problem in $\mathbf{R}^2$:

$$\begin{aligned}
\min_{x_1,x_2} \quad & e^{-x_2} \\
\text{s.t.} \quad & \|x\|_2 \leq x_1 \\
& x_2 \geq 0
\end{aligned}$$

Exercise 3.1 (Representation Issue) Through this example, we are going to see that duality gap is closely related to the "representation" of the constraints.

(a) Find the feasible set and optimal value.

(b) Let us write the problem as

$$\min\{e^{-x_2} : g(x) \le 0, x \in X\}$$

where $g(x) = \|x\|_2 - x_1$ and $X = \{(x_1, x_2) : x_2 \ge 0\}$. Does the slater condition holds? What is the dual optimal value? Is there a duality gap?

(c) Let us write the problem as

$$\min\{e^{-x_2} : g(x) \le 0, x \in X\}$$

where $g(x) = -x_2$ and $X = \{(x_1, x_2) : \|x\|_2 \le x_1\}$. Does the slater condition holds? What is the dual optimal value? Is there a duality gap?

Solution

(a) The feasible set is $\{(x_1, x_2) : x_1 \ge 0, x_2 = 0\}$. The optimal value is 1.

(b) The Slater condition does not hold because there exists no $x \in X$ such that $g(x) < 0$. The Lagrange function is

$$L(x_1, x_2, \lambda) = e^{-x_2} + \lambda(\sqrt{x_1^2 + x_2^2} - x_1)$$

For any $\lambda > 0$, the Lagrange dual function is

$$\underline{L}(\lambda) = \inf_{x \in X} L(x_1, x_2, \lambda) = \inf_{x_1 \in \mathbf{R}, x_2 \ge 0} e^{-x_2} + \lambda(\sqrt{x_1^2 + x_2^2} - x_1) = 0$$

Hence, the optimal value of the Lagrange dual is equal to 0. Therefore, there is a duality gap.

(c) The (relaxed) Slater condition holds because there is only an equality constraint. The Lagrange function is

$$L(x_1, x_2, \lambda) = e^{-x_2} + \lambda(-x_2)$$

For any $\lambda > 0$, the Lagrange dual function is

$$\underline{L}(\lambda) = \inf_{x \in X} L(x_1, x_2, \lambda) = \inf_{x_1, x_2 : \|x\| \le x_1} e^{-x_2} - \lambda x_2 = \inf_{x_1, x_2 = 0} e^{-x_2} - \lambda x_2 = 1$$

Hence, the optimal value of the Lagrange dual is equal to 1. Therefore, there is no duality gap.

Remark. From this example, we can see that the duality gap issue depends on the "representation" of the constraints.

Problem 4: Application in Finance

Consider assests $S_1, \dots, S_n (n \ge 2)$ with random returns $\xi_1, \dots \xi_n$. Let μ_i and σ_i denote the expected return and standard deviation of the random return of asset S_i, and ρ_{ij} denote the correlation coefficient of the returns of asset S_i and S_j. Denote $\mu = [\mu_1; \dots; \mu_m]$ as the expected return of all assests, i.e. $\mathbf{E}[\xi] = \mu$. Denote $\Sigma = (\sigma_{ij})$ as the covariance matrix of the asset returns with $\sigma_{ii} = \sigma_i^2$ and $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ for $i \ne j$, i.e. $\mathbf{Var}(\xi) = \Sigma$.

Suppose an investor plans to invest the proportion x_i of his total funds in asset i, for $i = 1, \dots, m$. The resulting portfolio is represented as $x = (x_1, \dots, x_n)$ and $\sum_{i=1}^n x_i = 1$. Assume short sale is not allowed, i.e. $x_i \ge 0$ for any i. The investor wants to find the best portfolio strategy to maximize his total expected return and meanwhile minimize his "risk". One way to take both criterions into account is to minimize a linear combination

$$-\mathbf{E}[\xi^T x] + \lambda \mathbf{Var}[\xi^T x]$$

where $\lambda > 0$ is a risk-aversion constant and balances the resturn and risk.

Exercise 4.1 Assume there are no restrictions on the portfolio. Formulate the above problem into an optimization model. Is the problem convex or not?

Solution The optimization problem can be modeled as :

$$\begin{aligned} \min_x \quad & -\mu^T x + \lambda \cdot x^T \Sigma x \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1, \ldots, n \end{aligned}$$

The problem is convex.

Exercise 4.2 Write down the Karush-Kuhn-Tucker optimality conditions for the problem.

Solution Let us reformulate the problem into the standard form:

$$\begin{aligned} \min_x \quad & -\mu^T x + \lambda \cdot x^T \Sigma x \\ & \sum_{i=1}^n x_i - 1 \leq 0 && (y_1) \\ & 1 - \sum_{i=1}^n x_i \leq 0 && (y_2) \\ & -x \leq 0 && (z) \end{aligned}$$

Let $y = (y_1; y_2) \geq 0, z \geq 0$ be the Lagrange multipliers. The Lagrange function is

$$L(x, y, z) = -\mu^T x + \lambda x^T \Sigma x + (y_1 - y_2)(\sum_{i=1}^n x_i - 1) - z^T x$$

Let x^* and (y^*, z^*) be the optimal primal-dual pair. The KKT condition says

$$\begin{aligned} (a) \quad & -\mu + 2\lambda \cdot \Sigma x^* + (y_1^* - y_2^*)\mathbf{1} - z = 0 \\ (b) \quad & y_1^*(\sum_{i=1}^n x_i^* - 1) = 0 \\ (c) \quad & y_2^*(1 - \sum_{i=1}^n x_i^*) = 0 \\ (d) \quad & z_i^* x_i^* = 0,\ i = 1, \ldots, n \end{aligned}$$

Exercise 4.3 An alternative risk measure is Value-at-Risk(VaR) developed by financial engineers at J.P. Morgan. Given a probability level $\alpha \in (0, 1)$, the α-VaR of a random variable η is defined as:

$$\mathbf{VaR}_\alpha(\eta) := \min\{\gamma : P(\eta \geq \gamma) \leq 1 - \alpha\}$$

Now change the objective to minimize the Value-at-Risk of the total return, i.e., $\mathbf{VaR}_\alpha(\xi^T x)$ with some $\alpha > 0.5$. Simplify the new optimization problem. Is the new problem convex or not? What if when ξ is a Gaussian random vector?

Solution To minimize the Value-at-Risk, we need to solve the optimization problem:

$$\begin{aligned} \min_x \quad & \mathbf{VaR}_\alpha(\xi^T x) \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1, \ldots, n \end{aligned}$$

Invoking the definition of Value-at-Risk, this can be simplified as

$$\begin{aligned}\min_{x,\gamma} \quad & \gamma \\ & P(\xi^T x \geq \gamma) \leq 1-\alpha \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1,\ldots,n\end{aligned}$$

The problem may or may not be convex because the constraint function $P(\xi^T x \geq \gamma)$ may not be convex.

When ξ is a Gaussian random variable, we have $\xi^T x \sim N(\mu^T x, x^T \Sigma x)$. Since

$$P(\xi^T x \geq \gamma) = P(\frac{\xi^T x - \mu^T x}{\sqrt{x^T \Sigma x}} \geq \frac{\gamma - \mu^T x}{\sqrt{x^T \Sigma x}}) = P(Z \geq \frac{\gamma - \mu^T x}{\sqrt{x^T \Sigma x}})$$

we have

$$P(\xi^T x \geq \gamma) \leq 1-\alpha \iff \frac{\gamma - \mu^T x}{\sqrt{x^T \Sigma x}} \geq \Phi^{-1}(\alpha)$$

The original problem can be simplified as

$$\begin{aligned}\min_{x,\gamma} \quad & \gamma \\ & -\mu^T x + \Phi^{-1}(\alpha)\sqrt{x^T \Sigma x} - \gamma \leq 0 \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1,\ldots,n\end{aligned}$$

Note that $\sqrt{x^T \Sigma x} = \|\Sigma^{1/2} x\|_2$ is convex. Since $\alpha > 0.5$, we have $\Phi^{-1}(\alpha) > 0$. Hence, the constraint function $\mu^T x + \Phi^{-1}(\alpha)\sqrt{x^T \Sigma x} - \gamma$ is convex. Thus, the optimization problem is convex.

Exercise 4.4 A well-known modification of the Value-at-Risk is conditional Value-at-Risk (CVaR), which takes into account of the magnitute of random variables beyond the VaR value. Given a probability level $\alpha \in (0,1)$, the α-CVaR of a random variable η is defined as: $\mathbf{CVaR}_\alpha(\eta) := \mathbf{E}[\eta | \eta \geq \mathbf{VaR}_\alpha(\eta)]$. It can be shown that

$$\mathbf{CVaR}_\alpha(\eta) = \min_{\gamma > 0} \{\gamma + \frac{1}{1-\alpha}\mathbf{E}[(\eta - \gamma)_+]\}$$

where $u_+ := \max(u, 0)$. Now change the objective to minimize the conditional Value-at-Risk of the total return, i.e. $\mathbf{CVaR}_\alpha(\xi^T x)$ with some $\alpha > 0$. Simplify the new optimization problem. Is the new problem convex or not?

Solution To minimize the conditional Value-at-Risk, we need to solve the optimization problem:

$$\begin{aligned}\min_{x} \quad & \mathbf{CVaR}_\alpha(\xi^T x) \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1,\ldots,n\end{aligned}$$

Invoking the definition of conditional Value-at-Risk, this can be simplified as

$$\begin{aligned}\min_{x,\gamma > 0} \quad & \gamma + \frac{1}{1-\alpha}\mathbf{E}[(\xi^T x - \gamma)_+] \\ & \sum_{i=1}^n x_i = 1 \\ & x_i \geq 0,\ i = 1,\ldots,n\end{aligned}$$

Note that the function $(\xi^T x - \gamma)_+$ is convex in (x, γ) for any ξ, so the expectation $\mathbf{E}[(\xi^T x - \gamma)_+]$ is also convex in (x, γ). Thus, the optimization problem is convex.