IE 521 Convex Optimization
Homework #3

your NAME here
your NetID here

February 27, 2019

Instructions.

• Homework is due **Friday, March 8, at 11:59pm**; no late homework accepted.
• Please use the provided \LaTeX file as a template.
• You can discuss with others, but please write your own solutions.
Problem 1: Portfolio Optimization

Consider assets $S_1, \ldots, S_n (n \geq 2)$ with random returns $\xi_1, \ldots, \xi_n$. Let $\mu_i$ and $\sigma_i$ denote the expected return and standard deviation of the random return of asset $S_i$, and $\rho_{ij}$ denote the correlation coefficient of the returns of asset $S_i$ and $S_j$. Denote $\mu = [\mu_1, \ldots, \mu_m]$ as the expected return of all assets, i.e. $E[\xi] = \mu$. Denote $\Sigma = (\sigma_{ij})$ as the covariance matrix of the asset returns with $\sigma_{ii} = \sigma^2_i$ and $\sigma_{ij} = \rho_{ij}\sigma_i\sigma_j$ for $i \neq j$, i.e. $\text{Var}(\xi) = \Sigma$. 

Suppose an investor plans to invest the proportion $x_i$ of his total funds in asset $S_i$ for $i = 1, \ldots, m$. The resulting portfolio is represented as $x = (x_1, \ldots, x_n)$ and $\sum_{i=1}^n x_i = 1$. Assume short sale is not allowed, i.e. $x_i \geq 0$ 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 criterion into account is to minimize a linear combination

$$\lambda \cdot \text{Var}[\xi^T x] - E[\xi^T x]$$

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

Exercise 1.1 Assume there are no restrictions on the portfolio. Formulate the above problem into an optimization model. Is the problem convex or not?

Solution. Put your solution here.

Exercise 1.2 Write down the Karush-Kuhn-Tucker optimality conditions for the problem.

Solution. Put your solution here.

Exercise 1.3 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 $\alpha$-VaR of a random variable $\eta$ is defined as:

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

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

Solution. Put your solution here.

Exercise 1.4 A well-known modification of the Value-at-Risk is conditional Value-at-Risk (CVaR), which takes into account of the magnitude of random variables beyond the VaR value. Given a probability level $\alpha \in (0, 1)$, the $\alpha$-CVaR of a random variable $\eta$ is defined as: $\text{CVaR}_\alpha(\eta) := E[\eta | \eta \geq \text{VaR}_\alpha(\eta)]$. It can be shown that

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

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

Solution. Put your solution here.
Problem 2: Support Vector Machine

Support vector machine (SVM) is a popular model in machine learning used for classification. Mathematically, given a training dataset of \( m \) points \((x_1, y_1), \ldots, (x_m, y_m)\) where \( x_i \in \mathbb{R}^n \) stands for the feature vector and \( y_i \in \{1, -1\} \) stands for two classes. The goal is to find two parallel hyperplanes represented by \((w, b)\) with maximal margin that separates the two classes of data, such that for class with \( y_i = 1 \), we have \( w^T x_i + b \geq 1 \) and for class with \( y_i = -1 \), we have \( w^T x_i + b \leq -1 \). Hence, we wish to satisfy \( y_i (w^T x_i + b) \geq 1 \) for \( i = 1, \ldots, m \).

If the data is not fully separable, we allow for small margin errors \( \epsilon_i > 0 \), \( i = 1, \ldots, m \), and we wish to also minimize these errors. This leads to solving the following optimization problem:

\[
\min_{w, b, \epsilon} \frac{1}{2} \|w\|^2 + C \cdot \sum_{i=1}^{m} \epsilon_i \quad \text{s.t.} \quad y_i (w^T x_i + b) \geq 1 - \epsilon_i, \quad i = 1, \ldots, m \quad (P)
\]

where the parameter \( C > 0 \) plays a role in controlling the relative importance of minimizing the norm of \( w \) (i.e., maximizing the margin) and minimizing the errors. Note that this problem is indeed a convex optimization problem.

Exercise 2.1 (Lagrange Duality) Let \( \alpha \geq 0 \) and \( \beta \geq 0 \) be the Lagrange multipliers associated with the two constraints. Show that the Lagrange dual problem of \((P)\) is given by the quadratic program:

\[
\max_{\alpha} \quad \sum_{i=1}^{m} \alpha_i - \frac{1}{2} \sum_{i,j=1}^{m} \alpha_i \alpha_j y_i y_j (x_i^T x_j) \\
\text{s.t.} \quad \sum_{i=1}^{m} \alpha_i y_i = 0 \\
0 \leq \alpha_i \leq C, \quad i = 1, \ldots, m \quad (D)
\]

Moreover, show that the primal and dual optimal solutions satisfy that

\[
\begin{align*}
\alpha_i = 0 & \Rightarrow y_i (w^T x_i + b) \geq 1 \\
\alpha_i = C & \Rightarrow y_i (w^T x_i + b) \leq 1 \\
0 < \alpha_i < C & \Rightarrow y_i (w^T x_i + b) = 1
\end{align*}
\]
We call the data points with non-zero Lagrangian multipliers the support vectors.

**Solution.** Put your solution here.

**Exercise 2.2** Show that \((P)\) can be equivalently written as an unconstrained convex problem

\[
\min_{w,b} \frac{1}{m} \sum_{i=1}^{m} \max(1 - y_i(w^T x_i + b), 0) + \lambda \|w\|_2^2 \quad (P')
\]

where \(\lambda > 0\) is some parameter.

**Solution.** Put your solution here.
Problem 3: Geometric Programming

Exercise 3.1  Consider the geometric program:

\[
\begin{align*}
\min_{x_1, \ldots, x_n} & \quad \sum_{j=1}^{J} c_j x_1^{a_j(1)} x_2^{a_j(2)} \cdots x_n^{a_j(n)} \\
\text{s.t.} & \quad \sum_{k=1}^{K} d_k x_1^{b_k(1)} x_2^{b_k(2)} \cdots x_n^{b_k(n)} \leq 1,
\end{align*}
\]

where \( c_j > 0, d_k > 0 \) and \( a_j \in \mathbb{R}^n, b_k \in \mathbb{R}^n \) for any \( j = 1, \ldots, J \) and \( k = 1, \ldots, K \). Note that the geometric program in general is not convex. An example would be

\[
\begin{align*}
\min_{x,y,z} & \quad \frac{x}{y} \\
\text{s.t.} & \quad x^2 + 3\sqrt{yz} \leq 1.
\end{align*}
\]

Show that the above general geometric program \((GP)\) can be reformulated into a convex program.

Solution. Put your solution here.

Exercise 3.2  Show that the Lagrange dual of the following convex program

\[
\begin{align*}
\min_{x=(x_1, \ldots, x_n) \in \mathbb{R}^n} & \quad \log \sum_{j=1}^{J} \exp(a_j^T x + \beta_j) \\
\text{s.t.} & \quad \log \sum_{k=1}^{K} \exp(b_k^T x + \gamma_k) \leq 0.
\end{align*}
\]

is equivalent to

\[
\begin{align*}
\max_{u,v} & \quad \beta^T u - \sum_{j=1}^{J} u_j \log u_j + \gamma^T v - \sum_{k=1}^{K} v_k \log \frac{v_k}{1^T v} \\
\text{s.t.} & \quad u \geq 0, v \geq 0, \\
& \quad 1^T u = 0 \\
& \quad A^T u + B^T v = 0.
\end{align*}
\]

Here the matrix \( A \in \mathbb{R}^{J \times n} \) and \( B \in \mathbb{R}^{K \times n} \), where the \( j \)-th row of matrix \( A \) corresponds to \( a_j \), and the \( k \)-th row of matrix \( B \) corresponds to \( b_k \).

Solution. Put your solution here.
Problem 4: Power Control (Bonus: 20 pts)

Consider a wireless network with $n$ logical transmitter/receiver pairs. Maximizing the transmit power can often be formulated as the following (nonconvex) problem:

$$\max_{x \in \mathbb{R}^n, r \in \mathbb{R}^n} \sum_{i=1}^{n} \log(r_i)$$

s.t. $$r_i = \log \left( 1 + \frac{a_i x_i}{\sum_{k \neq i} a_k x_k + \sigma^2} \right), \ i = 1, \ldots, n,$$

$$0 \leq x_i \leq C, \ i = 1, \ldots, n.$$

Here $a_i > 0$ stands for the channel gain coefficient of each transmitter, $r_i$ denotes the signal-to-interference ratio, for any $i = 1, \ldots, n$, and the variable $x_i$ stands for the power assigned to each user on the network.

Exercise 4.1  Construct an equivalent convex reformulation of the above problem.

Solution.  Put your solution here.