Midterm: Solutions

1. (20 points) *Risk Budgeting.* We are given a symmetric, $n \times n$ positive-definite matrix C, and a vector $\theta \in \mathbf{R}^n_{++}$, with $\mathbf{1}^T\theta = 1$, where $\mathbf{1}$ is the n-vector of ones.

   We consider a *risk budgeting* problem arising in financial optimization, which consists in finding a vector $x \in \mathbf{R}^n$ such that

   $$x > 0, \quad x_i(Cx)_i = \theta_i(x^T Cx), \quad i = 1, \ldots, n.$$

   (In case you are curious: the term "risk budgeting" refers to the fact that each so-called "partial risk" $x_i(Cx)_i$ is assigned a fixed proportion θ_i of the total risk, defined as the variance $x^T Cx$, itself the sum of the partial risks.)

   (a) Is the problem of finding a risk budgeting portfolio, as defined above, convex? Justify your answer carefully.

   (b) Define the function $h : \mathbf{R}^n \to \mathbf{R}$, with domain $\mathbf{R}^n_{++}$, and values for $x \in \mathbf{R}^n_{++}$ given by

   $$h(x) := -\sum_{i=1}^n \theta_i \log x_i.$$

   Now consider the problem

   $$p^* := \min_x \; h(x) \; : \; x^T Cx \leq 1. \tag{1}$$

   Show that the above problem is convex, strictly feasible, and attained.

   (c) Show that the KKT conditions for a primal-dual pair $(x, \lambda) \in \mathbf{R}^{n+1}$ are

   $$x > 0, \quad \frac{\theta_i}{x_i} - 2\lambda(Cx)_i = 0, \quad i = 1, \ldots, n, \quad \lambda(1 - x^T \Sigma x) = 0.$$

   Do these conditions charaterize optimality?

   (d) Show that $\lambda = 1/2$ is the (unique) optimal dual point. *Hint:* remember that $\mathbf{1}^T\theta = 1$.

   (e) Show that the risk budgeting problem always has a solution, which can be found by solving the convex problem (1).

   Solution:

   (a) The risk budgeting problem is not convex as given, since it involves quadratic equality constraints.

(b) The problem is convex, since h is convex, and the matrix C is positive-definite. It is strictly feasible: pick a vector $x \in \mathbf{R}^n_{++}$, and scale it so that the constraint is strictly satisfied.

(c) The Lagrangian writes

$$\mathcal{L}(x, \lambda) = -\sum_{i=1}^{n} \theta_i \log x_i + \lambda(x^T C x - 1).$$

The KKT conditions are:

- primal feasibility: $x > 0$ (so that x is in the domain of h);
- dual feasibility: $\lambda \geq 0$;
- Lagrangian stationarity:

$$\frac{\theta_i}{x_i} - 2\lambda(\Sigma x)_i = 0, \quad i = 1, \ldots, n;$$

- complementarity: $\lambda(1 - x^T C x) = 0$.

(d) The stationarity conditions imply that

$$\theta_i = 2\lambda x_i (Cx)_i, \quad i = 1, \ldots, n.$$

Summing, and using $\mathbf{1}^T \theta = 1$, leads to $1 = 2\lambda x^T C x$. The complementarity conditions then imply $\lambda = 1/2$.

(e) The KKT conditions characterize optimality since the primal problem is strictly feasible and convex. Those conditions are exactly the ones arising in the risk budgeting problem, hence the result.

2. (15 points) *Fenchel conjugate.*

For $\epsilon \geq 0$, $\lambda \geq 0$, consider the function with domain $\mathbf{R}^n$, and values for $w \in \mathbf{R}^n$ given by

$$f(w) = \epsilon \|w\|_2 + \lambda \|w\|_1.$$

In this exercise, we work out the Fenchel conjugate of f, which has values for $z \in \mathbf{R}^n$ given by

$$f^*(z) = \max_w \; z^T w - f(w).$$

(a) Set $\lambda = 0$. Show that the Fenchel conjugate has values for $z \in \mathbf{R}^n$ given by

$$f^*(z) = \begin{cases} 0 & \text{if } \|z\|_2 \leq \epsilon, \\ +\infty & \text{otherwise.} \end{cases}$$

(b) Derive a similar expression for the case when $\epsilon = 0$, precisely:

$$f^*(z) = \begin{cases} 0 & \text{if } \|z\|_\infty \leq \lambda, \\ +\infty & \text{otherwise.} \end{cases}$$

(c) Show that, in the general case:

$$f^*(z) = \begin{cases} 0 & \text{if there exist } \nu \in \mathbf{R}^n \text{ such that } \|\nu\|_\infty \leq \lambda, \ \ \|z-\nu\|_2 \leq \epsilon, \\ +\infty & \text{otherwise.} \end{cases}$$

Hint: express $f^*(z)$ as

$$f^*(z) = \max_{v,w} \ z^T w - \epsilon\|w\|_2 - \lambda\|v\|_1 \ : \ w = v,$$

then apply Lagrange duality and the results of the previous parts.

Solution:

(a) When $\lambda = 0$, we have

$$f^*(z) = \max_{w} \ z^T w - \epsilon\|w\|_2.$$

We note that $w = 0$ is feasible for the above problem, hence $f^*(z) \geq 0$. If $\|z\|_2 \leq \epsilon$, the Cauchy-Schwartz inequality implies that

$$\forall\, w, \ \ z^T w \leq \epsilon\|w\|_2,$$

hence $f^*(z) \leq 0$ and thus $f^*(z) = 0$. If $\|z\|_2 > \epsilon$, then $z \neq 0$, and we define $w(t) = tz/\|z\|_2$, with t in $\mathbf{R}_+$: we have then

$$z^T w(t) - \epsilon\|w(t)\|_2 = t(\|z\|_2 - \epsilon) \rightarrow +\infty \text{ when } t \rightarrow +\infty,$$

which proves the result.

(b) When $\epsilon = 0$, a similar reasoning pertains, due to the generalized Cauchy-Schwartz inequality: if $\|z\|_\infty \leq \lambda$, then

$$\forall\, w, \ \ z^T w \leq \lambda\|w\|_1.$$

(c) The problem

$$f^*(z) = \max_{v,w} \ z^T w - \epsilon\|w\|_2 - \lambda\|v\|_1 \ : \ w = v,$$

is convex and strictly feasible. Hence, strong duality holds. Applying Lagrange duality, we have

$$\begin{aligned} f^*(z) &= \max_{v,w} \ \min_{\nu} \ z^T w - \epsilon\|w\|_2 - \lambda\|v\|_1 + \nu^T(w - v) \\ &= \min_{\nu} \ \max_{v,w} \ z^T w - \epsilon\|w\|_2 - \lambda\|v\|_1 + \nu^T(w - v). \end{aligned}$$

The maximizations over w, v can be done separately, using the previous derivations. We obtain

$$f^*(z) = \min_{\nu} \ 0 \ : \ \|\nu\|_\infty \leq \lambda, \ \ \|z - \nu\|_2 \leq \epsilon,$$

which is the desired result, since in a minimization problem, by convention, the optimal value is $+\infty$ iff the problem is not feasibe.

3. (25 points) *Square-root elastic net.* Consider the problem

$$p^* = \min_w \|X^T w - y\|_2 + \epsilon \|w\|_2 + \lambda \|w\|_1,$$

where $X \in \mathbf{R}^{n \times m}$, $y \in \mathbf{R}^m$, $\epsilon > 0$ and $\lambda > 0$ are given.

(a) Show that a dual can be written as

$$\max_{\nu, v} \nu^T y \;:\; \|X\nu - v\|_\infty \leq \lambda, \;\; \|v\|_2 \leq \epsilon, \;\; \|\nu\|_2 \leq 1.$$

Hint: Introduce a new variable $r := X^T w - y$, then apply Lagrange duality, and the result of problem 2; alternatively, you can use Sion's theorem and conic duality.

(b) Denote by a_i the i-th column of X^T. Show that if

$$\|a_i\|_2 < \lambda - \epsilon$$

then we can set $w_i = 0$ at the optimum of the primal problem. *Hint:* show that for a given vector $a \in \mathbf{R}^m$, and scalar c, we have

$$\max_{u \,:\, \|u\|_2 \leq 1} |a^T u - c| \leq \|a\|_2 + |c|,$$

and use the fact that, in a convex problem, a constraint that is inactive at optimum can be safely removed.

(c) Prove the following statements.

i. p^*, considered as a function of y, is convex.

ii. p^*, considered as a function of (ϵ, λ) with domain $\mathbf{R}^2_+$, is concave.

iii. p^*, considered as a function of (ϵ, λ) with domain $\mathbf{R}^2_+$, is not convex in general. *Hint:* consider the case when X, y are both scalars and equal to 1, and $\lambda = 0$; show graphically that then the optimal w is either 0 or 1.

Solution:

(a) We have, with $f(w) := \epsilon \|w\|_2 + \lambda \|w\|_1$:

$$\begin{aligned} p^* &= \min_{r,w} \|v\|_2 + f(w) \;:\; r = X^T w - y \\ &= \min_{r,w} \max_{\nu} \|u\|_2 + f(w) + \nu^T (r - X^T w + y) \\ &= \max_{\nu} \min_{r,w} \|v\|_2 + f(w) + \nu^T (r - X^T w + y), \end{aligned}$$

where the last line derives from Slater's strong duality result. We observe that the dual function involves problems that can be solved in v, w separately. We further obtain

$$\begin{aligned} p^* &= \max_{\nu} \nu^T y + \min_{v} \|v\|_2 + \nu^T v + \min_{w} f(w) - w^T X \nu \\ &= \max_{\nu} -f^*(-X\nu) \;:\; \|\nu\|_2 \leq 1. \end{aligned}$$

According to the result of part 2, The value $f^*(-X\nu)$ is finite (and zero) if and only if there exist v such that

$$\|-X\nu - v\|_\infty \leq \lambda, \;\; \|v\|_2 \leq \epsilon.$$

Changing v into $-v$ is harmless, and the desired result follows.
Alternatively, we can use the conic Lagrangian

$$\mathcal{L}(w,p,q,r) = u^T(y - X^T w) + v^T w + r^T w,$$

so that, defining the compact set

$$\mathcal{Z} := \{(u,v,r) \;:\; \|u\|_2 \leq 1, \;\; \|v\|_2 \leq \epsilon, \;\; \|r\|_\infty \leq \lambda\},$$

we obtain the min-max representation:

$$p^* = \min_w \max_{(u,v,r)\in\mathcal{Z}} \mathcal{L}(w,u,v,r).$$

Using Sion's theorem, we obtain

$$p^* = \max_{(u,v,r)\in\mathcal{Z}} \min_w \mathcal{L}(w,u,v,r),$$

Solving for w, we get

$$p^* = \max_{(u,v,r)\in\mathcal{Z}} u^T y \;:\; Xu = v + r.$$

Eliminating the variable r leads to the desired result.

(b) The constraint $\|Xu - v\|_\infty \leq \lambda$ expresses as

$$|a_i^T u - v_i| \leq \lambda, \;\; i = 1, \ldots, n.$$

For a given vector $a \in \mathbf{R}^n$, and scalar c, we have

$$\max_{u \,:\, \|u\|_2 \leq 1} |a^T u - c| \leq \max_{u \,:\, \|u\|_2 \leq 1} |a^T u| + |c| = \|a\|_2 + |c|.$$

(Note that equality holds, which is seen upon choosing $u = \mathbf{sign}(c) a / \|a\|_2$.) This shows that for a given index $i \in \{1, \ldots, n\}$, we have

$$\max_{u,v} \left\{ |a_i^T u - v_i| \;:\; \|v\|_2 \leq \epsilon, \;\; \|u\|_2 \leq 1 \right\} \leq \|a_i\|_2 + \max_{v \,:\, \|v\|_2 \leq \epsilon} |v_i| = \|a_i\|_2 + \epsilon.$$

Assume now that the condition $\|a_i\|_2 < \lambda - \epsilon$ is met for some $i \in \{1, \ldots, n\}$. Then the constraint $|a_i^T u - v_i| \leq \lambda$ cannot be active at optimum; the i-th constraint involving a_i can thus be removed from the dual problem. This means that the value of the problem is the same when we replace a_i by zero in the primal; this is the same as setting $w_i = 0$.

(c) i. This is true, since the objective function is jointly convex in w, y.

ii. This is true, since the optimal value is a pointwise minimum of affine functions in (ϵ, λ).

iii. This is not true, since the function is concave, and not affine. This can be seen by taking $\lambda = 0$, $X = y = 1$, in which case the problem reads

$$p^* = \min_w \ |w - 1| + \epsilon|w|.$$

The optimum is attained at $w = 0$ or $w = 1$, so that

$$p^* = \min(1, \epsilon),$$

which is not affine.