EE 127 Final

NAME: **SID:**

The exam lasts 3 hours. The maximum number of points is 40. Notes are not allowed, except for a two-sided cheat sheet of regular format.

This booklet is 12 pages total, with extra blank spaces allotted throughout, and 2 blank pages at the end, left for you to write your answers.

There are three separate problems. All the questions in this exam can be solved independently of each other.

1. (*10 points*) Let $X \in \mathbf{R}^{n\times m}$, $y \in \mathbf{R}^m$ and $\lambda, \mu > 0$. Consider the so-called (square-root) "elastic net" problem

$$\min_w \|X^T w - y\|_2 + \lambda\|w\|_1 + \mu\|w\|_2. \tag{1}$$

   (a) Let $k << \min(m,n)$ be given. Assume that we know the top k singular values σ_i, $i = 1,\ldots,k$, of X, and the associated left and right singular vectors, denoted $u_i \in \mathbf{R}^n$, $v_i \in \mathbf{R}^m$, $i = 1,\ldots,k$. Construct matrices $L \in \mathbf{R}^{n\times k}$, $R \in \mathbf{R}^{n\times k}$ such that $X \approx LR^T$, with $R^T R = I_k$ (the $k \times k$ identity matrix).

   (b) Replace X by its low-rank approximation in problem (1). Show that the new problem reduces to one with n variables and k measurements (that is, the new vector y has dimension k).

   (c) In practice, it may be advisable to take into account the size of the error made in approximating X. How would you modify the problem to account for this error? *Hint:* you may assume the singular value σ_{k+1} is known, and use the fact that, for any $n \times m$ matrix Z, m-vector y and scalar $\epsilon > 0$:

$$\max_{\Delta\,:\,\|\Delta\|\le\epsilon} \|(Z+\Delta)^T w - y\|_2 = \|Z^T w - y\|_2 + \epsilon\|w\|_2,$$

   with $\|\Delta\|$ the largest singular value of matrix Δ.

   (d) The computational complexity of problem (1) grows in $O(nm^2 + m^3)$. What is the complexity estimate of the problem you obtained in part 1c? (If you are not sure about that part, use the formulation you found in part 1b.)

2. (*10 points*) The log-sump-exp function is defined as the function $f : \mathbf{R}^n \to \mathbf{R}$ with values

$$f(z) = \log\left(\sum_{i=1}^{n} e^{z_i}\right).$$

Often, the convexity of this function is proven by deriving the Hessian. In this exercise, you will show that f is convex, using another method.

(a) Show that, for any given $s > 0$, we have

$$\log s = -1 + \min_{v}\ se^{v} - v.$$

(b) Show that

$$f(z) = -1 + \min_{v} \sum_{i=1}^{n} e^{z_i+v} - v.$$

(c) Prove convexity of f based on the above result. *Hint:* use the notion of joint convexity.

3. (*10 points*) *Fermat point on a triangle:* Consider a triangle with distinct vertices A_1, A_2, A_3 in two dimensions. We seek the point P that minimizes the sum of the distances to the vertices of the triangle. Representing the points as three distinct vectors $a_i \in \mathbf{R}^2$, $i = 1, 2, 3$, and the (unknown) location of P as $x \in \mathbf{R}^2$, the problem writes as

$$p^* = \min_x \sum_{i=1}^{3} \|x - a_i\|_2.$$

In this exercise, we show that the point P satisfies $\widehat{A_1 P A_2} = \widehat{A_2 P A_3} = \widehat{A_3 P A_1} = 120°$.

(a) Use Sion's theorem to show that the problem admits the dual expression:

$$p^* = \max_{y_1, y_2, y_3} \sum_{i=1}^{3} a_i^T y_i \ : \ y_1 + y_2 + y_3 = 0, \quad \|y_i\|_2 \le 1, \quad i = 1, 2, 3.$$

(b) Assume that the primal optimal point x^* is unique, and that $x^* \neq a_i$ for every $i = 1, 2, 3$. Show that optimal dual variables y_i^*, $i = 1, 2, 3$ satisfy

$$y_i^* = \frac{a_i - x^*}{\|x^* - a_i\|_2}, \quad i = 1, 2, 3.$$

Hint: show that strong duality implies that

$$\sum_{i=1}^{3} (a_i - x^*)^T y_i^* = \sum_{i=1}^{3} \|x^* - a_i\|_2.$$

and reason by contradiction, invoking Cauchy-Schwartz.

(c) Show that the dual solution vectors satisfy $y_i^T y_j = -1/2$, $i \neq j$. *Hint:* Express $\|y_1\|_2^2$ as a function of y_2, y_3.

(d) Conclude the argument.

4. (*10 points*) We consider an investment problem of the form

$$p^* = \max_{x \geq 0} \; r^T x - \frac{1}{2} x^T C x$$

where $r \in \mathbf{R}^n$ is a vector of expected returns, $C = C^T \succ 0$ is a covariance matrix. We assume that C is given as a so-called "single factor model", that is

$$C = D + f f^T,$$

where D is diagonal positive-definite, and $f \in \mathbf{R}^n$.

(a) First show a preliminary result: for every scalars ρ and $\delta > 0$,

$$\max_{\xi \geq 0} \; \rho \xi - \frac{1}{2} \delta \xi^2 = \frac{1}{2\delta} \max(0, \rho)^2.$$

with unique minimizer $\xi^* = \max(0, \rho)/\delta$.

(b) Based on the expression

$$p^* = \max_{x \geq 0} \; r^T x - \frac{1}{2} x^T D x - \frac{1}{2} z^2 \;\; : \;\; z = f^T x,$$

show that a dual can be written as

$$\frac{1}{2} \min_{\nu} \; \nu^2 + \sum_{i=1}^{n} \frac{\max(0, r_i - \nu f_i)^2}{D_{ii}}.$$

Hint: for the minimization of the Lagrangian over x, use part 4a .

(c) Does strong duality hold? Justify your answer.

(d) Explain how to recover an optimal primal dual point from an optimal dual variable ν^*.

EXTRA SPACE

EXTRA SPACE

EXTRA SPACE