OR 6327: Semidefinite Programming. Spring 2012.

Takehome Final Exam. Pickup Tuesday, May 15th, 9 am (also available electronically); due 12 m (noon) Wednesday.

This is to be your work solely. You can refer to the class notes and to any results from homeworks, but no other sources (living, dead tree, or digital). You can also use standard results in algebra, analysis, and convex analysis, and of course, you can use an earlier part of a question in a later part, whether you have completed the first part or not. If you need a hint, contact me in my office (229 Rhodes), by email (mjt7), or by phone (255-9135 office, 257-3344 home, before 10:30 pm).

1. a) Consider the standard form primal "knapsack" SDP

$$\min C \bullet X, \quad A \bullet X = \beta, \quad X \succeq 0,$$

with a single equality constraint, where $C \in \mathbb{M}^n$, $A \in \mathbb{M}^n_{++}$, and $\beta > 0$. Show how to find an optimal solution to this problem and to its dual using one Cholesky factorization and one eigenvalue decomposition.

b) We showed that $-\mathrm{lndet}(X)$ is convex on the set of positive definite matrices, in analogy with $-\ln \xi$ on the positive real line. In addition, ξ^{-1} is convex on the positive real line. Show that $f(X) := X^{-1} \bullet W$ is convex on the set of positive definite matrices for any fixed $W \succ 0$.

2. The spectral bundle method is designed to solve the following SDP and its dual:

$$\begin{array}{lrcl} \max_X & C \bullet X & & \\ (P) & \mathcal{A}X & = & 0, \\ & I \bullet X & = & 1, \\ & X & \succeq & 0, \end{array} \qquad \begin{array}{lrcl} \min_{\lambda,y} & \lambda & & \\ (D) & \lambda I + \mathcal{A}^* y & \succeq & C, \end{array}$$

where the dual can also be written $\min_y g(y) := \lambda_{\max}(C - \mathcal{A}^* y)$.

Suppose you have produced a number of subgradients $-\mathcal{A}(q_i q_i^T)$ ($q_i \in \mathbf{R}^n$, $\|q_i\|_2 = 1$) of g at y and nearby points, and for all i, $q_i^T (C - \mathcal{A}^* y) q_i \geq \lambda_{\max}(C - \mathcal{A}^* y) - \epsilon$ for some $\epsilon > 0$. Suppose also that you know a convex combination of these subgradients that equals 0. Show how to obtain ϵ-optimal solutions to (P) and (D).

3. a) Except for solutions where $\tilde{S}$ does not have full rank, the central path for a standard form SDP and its dual can be defined by the primal and dual feasibility conditions and the equation $\tilde{S}\tilde{X}\tilde{S} = \mu \tilde{S}$. (Note that whenever $\tilde{X}$ and $\tilde{S}$ are symmetric, so is $\tilde{S}\tilde{X}\tilde{S} - \mu \tilde{S}$.) Suppose you have a strictly feasible (X, y, S). Write down the Newton equations to approximate the (limit) point on the path for $\mu = 0$ using this equation as the third one defining the central path.

b) Say the solution to the system in (a) is $(\Delta X, \Delta y, \Delta S)$. Can you relate this to one of the directions we studied in the class?

c) What goes wrong with your answer in (b) if $X \succ 0$, $S \succ 0$, but (y, S) is not feasible in (D)?

OVER

4. a) Suppose that $C \in \mathbb{M}^k_{++}$ and $M := \begin{bmatrix} A & B \\ B^T & C \end{bmatrix} \in \mathbb{M}^{n+k}_{++}$, with

$$M^{-1} = \begin{bmatrix} U & V \\ V^T & W \end{bmatrix},$$

with $U \in \mathbb{M}^n$. Show that $A - BC^{-1}B^T$ is invertible, with inverse U.

b) Assume we have a large set of random variables with a joint Gaussian distribution, but that we expect most pairs to be conditionally independent. We have seen that conditional independence corresponds to zeros in the inverse of the covariance matrix. However, suppose that of these variables, only n are observed, while the other k ($k \ll n$) are hidden and cannot be observed. We take a large number of samples of the observable variables, and assume for simplicity we have the exact covariance matrix, say U, of these variables. Why would you expect the inverse of U to be the sum of a low-rank matrix and a sparse matrix?

c) Using the material of the last part of the course, formulate a convex optimization problem to try to determine a decomposition of a given matrix $D \in \mathbb{M}^n$ into the sum of a low-rank symmetric matrix and a sparse symmetric matrix.

d) Explain *briefly* why

$$\begin{array}{lccccc} \min_{L,S} & \|L\|_* & + & |||S|||_1 & & \\ & L & + & S & = & D \\ & L \in \mathbb{M}^n, & & S \in \mathbb{M}^n, & & \end{array}$$

can be transformed into an SDP. However, in the large-scale case, it is likely to be more efficient to treat it directly via a first-order method. Again based on the last part of the course, suggest such a method whose steps can be carried out efficiently. (Details can be referred to the course notes.)

Have a great summer!