Midsem 2013 Solutions

30 Marks, Open Notes, 2 Hours. I have made every effort to ensure that all required assumptions have been stated. If absolutely necessary, do make more assumptions and state them very clearly.

1. In the class, we gave an analytic proof for the strong duality theorem for Linear Programs. In this question, we will attempt to give a geometrically motivated proof (for a part of the strong duality theorem) and your task will be to provide rigourous proofs for claims made in the process.

   Let A be an $m \times n$ matrix of reals, that is, $A \in \Re^{m \times n}$. Let P be the primal linear program given in (1)

$$\begin{aligned} \min_{\mathbf{x} \in \Re^n} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \geq \mathbf{b} \end{aligned} \tag{1}$$

   and D be the dual program given in (2)

$$\begin{aligned} \max_{\mathbf{y} \in \Re^m} \quad & \mathbf{b}^T\mathbf{y} \\ \text{subject to} \quad & A^T\mathbf{y} = \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{aligned} \tag{2}$$

   Suppose $\mathbf{x}^*$ is an optimal feasible solution for P . Let $\mathbf{a}_i^T\mathbf{x}^* \geq b_i$ for all $i \in I$ be all the constraints tight[1] at $\mathbf{x}^*$. Here, vector $\mathbf{a}_i^T$ is the i^{th} row of A. In other words, I is the index of all inequalities in the primal, that become equalities at $\mathbf{x}^*$.

   (a) We claim that the objective function vector $\mathbf{c}$ is contained in the cone $K = \left\{ \mathbf{x} \;\middle|\; \mathbf{x} = \sum_{i \in I} \lambda_i \mathbf{a}_i, \lambda_i \geq 0 \right\}$ generated by the set of vectors

[1] A constraint $\mathbf{a}_i^T\mathbf{x} \geq b_i$ is tight if $\mathbf{a}_i^T\mathbf{x} = b_i$

$\{\mathbf{a}_i\}_{i\in I}$. Prove this claim. Hints: prove by contradiction and make use of the separating hyperplane theorem.

ANS: Suppose for contradiction that $\mathbf{c}$ does not lie in this cone. Then there must exist a separating hyperplane between $\mathbf{c}$ and K: *i.e.*, there exists a vector $\mathbf{d} \in \Re^n$ such that $\mathbf{a}_i^T\mathbf{d} \geq 0$ for all $i \in I$, but $\mathbf{c}^T\mathbf{d} < \mathbf{0}$. Now consider the point $\mathbf{z} = \mathbf{x}^* + \epsilon\mathbf{d}$ for some tiny $\epsilon > 0$. Note the following:

i. For small enough ϵ, the point $\mathbf{z}$ satisifes the constraints $A\mathbf{z} \geq \mathbf{b}$. We prove this as follows.
   For $j \in I$, we have $\mathbf{a}_j^T\mathbf{z} = \mathbf{a}_j^T\mathbf{x}^* + \epsilon\mathbf{a}_j^T\mathbf{d} = b_j + \epsilon\mathbf{a}_j^T\mathbf{d} \geq b_j$ since $\epsilon > 0$ and $\mathbf{a}_j^T\mathbf{d} \geq 0$.
   For $j \notin I$, by choosing small enough $\epsilon \leq \min_j \left(\frac{b_j - \mathbf{a}_j^T\mathbf{x}^*}{\mathbf{a}_j^T\mathbf{x}^*}\right)$, we have $\mathbf{a}_j^T\mathbf{z} = \mathbf{a}_j^T\mathbf{x}^* + \epsilon\mathbf{a}_j^T\mathbf{d} \geq b_j$.

ii. The objective function value decreases since $\mathbf{c}^T\mathbf{z} = \mathbf{c}^T\mathbf{x}^* + \mathbf{c}^T\mathbf{d} < \mathbf{c}^T\mathbf{x}^*$.

This contradicts the fact that $\mathbf{x}^*$ was optimal.

(b) Therefore, the vector $\mathbf{c}$ lies within the cone $K = \left\{\mathbf{x} \;\middle|\; \mathbf{x} = \sum_{i\in I} \lambda_i\mathbf{a}_i, \lambda_i \geq 0\right\}$ generated by the set of vectors $\{\mathbf{a}_i\}_{i\in I}$. Present a choice of λ_i such that $\sum_i \lambda_i b_i = \mathbf{c}^T\mathbf{x}^*$

ANS: Choose λ_i for $i \in I$ so that $\mathbf{c} = \sum_{i\in I} \lambda_i\mathbf{a}_i$, $\lambda \geq 0$ and set $\lambda_j = 0$ for $j \notin I$.

- We know $\lambda \geq 0$.
- $A^T\lambda = \sum_{i\in I} \lambda_i\mathbf{a}_i = \mathbf{c}$
- $\mathbf{b}^T\lambda = \sum_{i\in I} b_i\lambda_i = \sum_{i\in I}(\mathbf{a}_i^T\mathbf{x}^*)\lambda_i = \mathbf{c}^T\mathbf{x}^*$

Therefore λ is a solution to the dual which yields dual objective value equal to that of primal.

(c) Prove that, if the primal has an optimal feasible solution, the dual must have an optimal feasible solution and that the optimal value of the objective for the dual equals the optimal value of the objective for the primal.

ANS: From the weak duality theorem, we know that the dual optimal cannot exceed the primal optimal. Since, for a given primal optimal, we have found a dual optimal that yields objective value equal to the primal optimal, we can be assured that the λ obtained above is a point of dual optimal.

(10 Marks)

2. We call S a *copositive* matrix if it is symmetric and satisfies $\mathbf{x}^T S\mathbf{x} \geq 0$ for all $\mathbf{x} \geq \mathbf{0}$. The set of *copositive* matrices is denoted by $\mathcal{S}_*^n$. In other words

$$\mathcal{S}_*^n = \left\{ S \,\middle|\, S \in \mathcal{S}^n,\ \mathbf{x}^T S\mathbf{x} \geq 0 \ \forall\ \mathbf{x} \geq \mathbf{0} \right\}$$

Is the set $\mathcal{S}_*^n$ of *copositive* matrices convex? Is $\mathcal{S}_*^n$ affine? Is $\mathcal{S}_*^n$ a convex cone? Is $\mathcal{S}_*^n$ a proper cone? Prove your claims.

ANS: $\mathcal{S}_*^n$ is a proper cone (and thefore a convex cone) for the following reason: it is obvious that all positive linear combinations of copositive matrices are also copositive. Hence $\mathcal{S}_*^n$ is a convex cone. $\mathcal{S}_*^n$ has a non-empty interior, because it includes the proper cone of positive semi-definite matrices (see page 5 of `http://www.cse.iitb.ac.in/~CS709/notes/eNotes/7-21-08-2013.pdf`). $\mathcal{S}_*^n$ is pointed because $S \in \mathcal{S}_*^n$ and $-S \in \mathcal{S}_*^n$ implies that $\mathbf{x}^T S\mathbf{x} = 0$ for all $\mathbf{x} \geq \mathbf{0}$ which implies that $S = 0$. It is not affine since the 2 by 2 matrices [5 0; 0 2] and [2 0; 0 3] are copositive BUT [5 0; 0 2] - [2 0; 0 3] is not copositive.

Let $A \in \mathcal{S}_*^n$. Find the dual cone of $\{A\mathbf{x} \mid \mathbf{x} \geq \mathbf{0}\}$, assuming A is a matrix of reals, $\mathbf{x}$ is a vector of reals and dot product in $\Re^n$ to be the inner product.

ANS: dual cone $= \left\{\mathbf{y} \mid A^T\mathbf{y} \geq \mathbf{0}\right\}$

(**8 Marks**)

3. Prove that the trace of a matrix equals the sum of its eigenvalues and that the determinant equals the product of its eigenvalues. You can use the fact that given $P(x) = \sum_{i=0}^{n} \alpha_i x^i$,

   (a) the sum of the roots of the polynomial $P(x)$ is always $\frac{-\alpha_{(n-1)}}{\alpha_n}$

   (b) the product of the roots of the polynomial $P(x)$ is always $(-1)^n \frac{\alpha_0}{\alpha_n}$

ANS: Eigenvalues are roots of the characteristic polynomial[2]. You only need $\alpha_0, \alpha_{(n-1)}$ and α_n.

(**5 Marks**)

4. The optimization problem in (3) is an *Integer Linear Program* (ILP).

$$\begin{aligned} \text{minimize} \quad & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & A\mathbf{x} \leq \mathbf{b} \\ & x_i \in \{0,1\} \text{ for } 1 \leq i \leq n \end{aligned} \tag{3}$$

In a general method called relaxation, the constraint that x_i be zero or one is replaced with the linear inequalities $0 \leq x_i \leq 1$. The problem in (4) is called the *Relaxation of the Linear Program* (RLP).

[2] `http://en.wikipedia.org/wiki/Characteristic_polynomial`

$$\begin{array}{lll} \text{minimize} & f(\mathbf{x}) = \mathbf{c}^T\mathbf{x} & \\ \text{subject to} & A\mathbf{x} \leq \mathbf{b} & (4) \\ & 0 \leq x_i \leq 1 \text{ for } 1 \leq i \leq n & \end{array}$$

It turns out that the RLP (4) is far easier to solve than the original ILP (3).

(a) What inequality relationship exists between the optimal value of the RLP (4) and the optimal value of the original ILP (3).

ANS: RLP $\leq$ ILP

(1 Mark)

(b) What can you say about the original ILP (3) if the RLP (4) is infeasible?

ANS: ILP is also infeasible.

(1 Mark)

(c) It sometimes happens that the RLP (4) has a solution with $x_i \in \{0, 1\}$ for all $1 \leq i \leq n$. What can you say in this case?

ANS: ILP will have the same solution as RLP.

(1 Mark)

5. Consider the half-space C and hyperbolic set D described below:

$$C = \left\{\mathbf{x} \in \Re^2 | x_2 \leq 0\right\}$$

and

$$D = \left\{\mathbf{x} \in \Re^2 | x_1x_2 \geq 1, x_1 \geq 0, x_2 \geq 0\right\}$$

Can C and D be separated by a hyperplane? Prove. State whether they can be strictly separated (no proof required).

ANS: Yes. They can be separated. All you need to prove is that each of them is convex and that they are disjoint and hence the separating hyperplane theorem can be applied. C is a half-space and obviously convex. For proving convexity of D, consider two points (x_1, x_2) and (y_1, y_2) in D. You need to show that their convex cominations also lie in D. Consider two cases: (i) $(x_1 - y_1)(x_2 - y_2) \geq 0$ and $(x_1 - y_1)(x_2 - y_2) < 0$. In each case, you just need to show that $(\theta x_1 + (1-\theta)y_1)(\theta x_2 + (1-\theta)y_2) \geq 1$.

They cannot be strictly separated.

(**4 Marks**)

6. Write the dual for the following (conic linear) program as compactly as possible:

$$\begin{aligned} \min_{x_1,x_2,x_3} \quad & 2x_1 + x_2 + x_3 \\ \text{subject to} \quad & \sqrt{x_2^2 + x_3^2} \leq x_1 \end{aligned} \tag{5}$$

(**5 Bonus Marks**)

ANS: You can write the constraint as a conic constraint in terms of a second order cone and use the fact that the dual cone of the second order cone is itself (the second order cone is self-dual). See page 2 of `http://www.cse.iitb.ac.in/~CS709/notes/eNotes/9-28-08-2013.pdf`. Here, $A = I_{3\times 3}$, $\mathbf{b} = \mathbf{0}$ and $K = \left\{x_1, x_2, x_3 | \sqrt{x_2^2 + x_3^2} \leq x_1\right\}$. However, the problem is pretty trivial. Objective in the dual is simply 0 – taking cue from the dual of the LP, which has objective in terms of $\mathbf{b}^T\mathbf{y}$, we note that the $\mathbf{b}$ vector here is just $\mathbf{0}$. Hence, the objective of the dual for the linear conic program will have 0, subject to conic constraints, which will not matter in the compact representation, other than for finding a feasible solution.

$$\begin{aligned} \max_{z_1,z_2,w} \quad & 0 \\ \text{subject to} \quad & ||\mathbf{z}|| \leq w \\ & z_2 = 1 \\ & z_3 = 1 \end{aligned} \tag{6}$$