Midsem 2015

38 Marks, 25% weightage, Open Notes, 2.5 Hours. You can assume anything that was stated in class. I have made every effort to ensure that all required additional assumptions have been stated. If absolutely necessary, do make more assumptions and state them very clearly.

1. A saddle point of a function $f : \mathcal{X} \times \mathcal{Y} \to \Re \cup \{\pm \inf\}$ is a pair $(\overline{x}, \overline{y}) \in \mathcal{X} \times \mathcal{Y}$ satisfying

$$\sup_{y} f(\overline{x}, y) \leq f(\overline{x}, \overline{y}) \leq \inf_{x} f(x, \overline{y})$$

Show that if $f(x, y)$ has a saddle point $(\overline{x}, \overline{y})$ then

$$\sup_{y} \inf_{x} f(x, y) = \inf_{x} \sup_{y} f(x, y)$$

(4 Marks)

Soln: By the min-max / inf-sup inequality proved in class:

$$\sup_{y} \inf_{x} f(x,y) \leq \inf_{x} \sup_{y} f(x,y) \rightarrow (1)$$

Now, if f has a saddle point $(\bar{x}, \bar{y})$ then

By defn of saddle pt

$$\inf_{x} \sup_{y} f(x,y) \leq \sup_{y} f(\bar{x}, y) \leq f(\bar{x}, \bar{y}) \leq \inf_{x} f(x, \bar{y}) \leq \sup_{y} \inf_{x} f(x,y)$$

Thus

$$\inf_x \sup_y f(x,y) \le \sup_y \inf_x f(x,y) \rightarrow 2$$

By 1 & 2, we have min-max equality!

$$\sup_y \inf_x f(x,y) = \inf_x \sup_y f(x,y) = f(\bar{x},\bar{y})$$

Illustration of saddle point at $(0,0)$ for $f(x_1,x_2) = x_1^2 - x_2^2$ on pages 242 & 243 of http://www.cse.iitb.ac.in/~cs709/notes/BasicsOfConvexOptimization.pdf

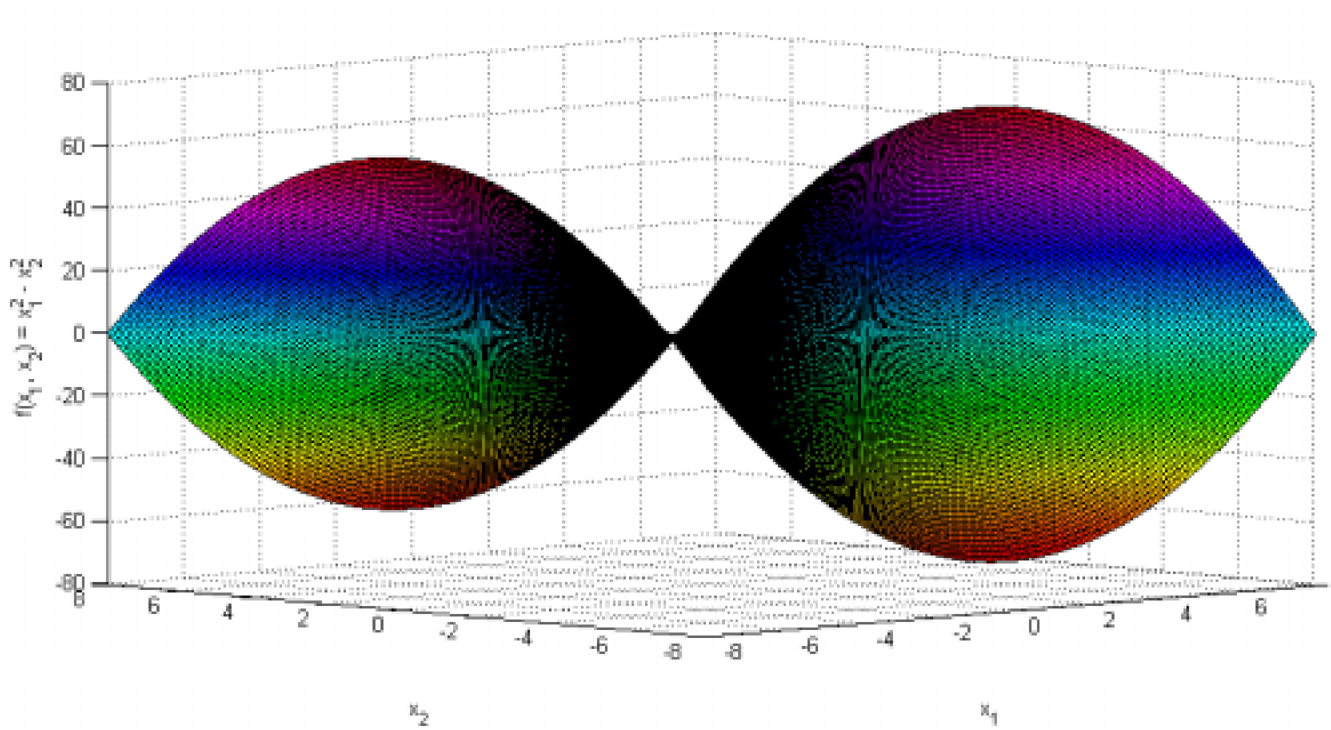

Figure 4.19: The hyperbolic paraboloid $f(x_1,x_2) = x_1^2 - x_2^2$, which l

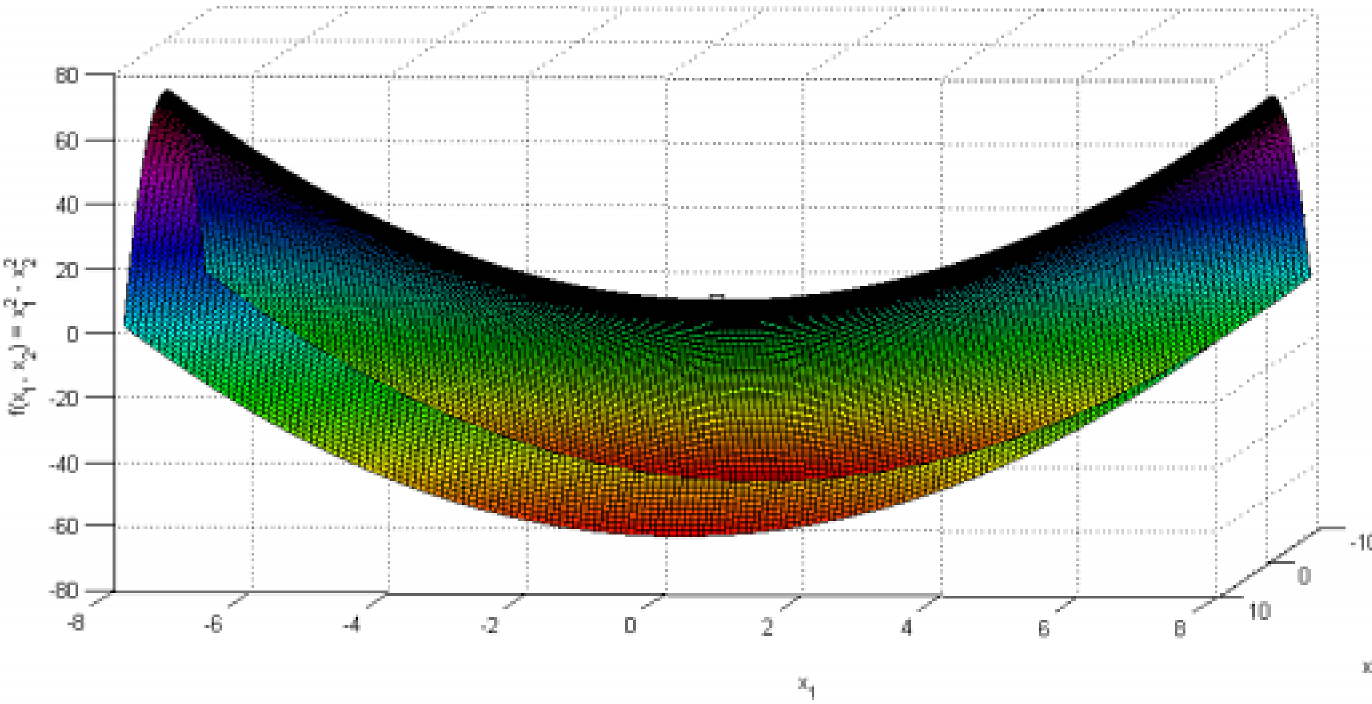

Figure 4.20: The hyperbolic paraboloid $f(x_1, x_2) = x_1^2 - x_2^2$, when vi
the x_1 axis is concave up.

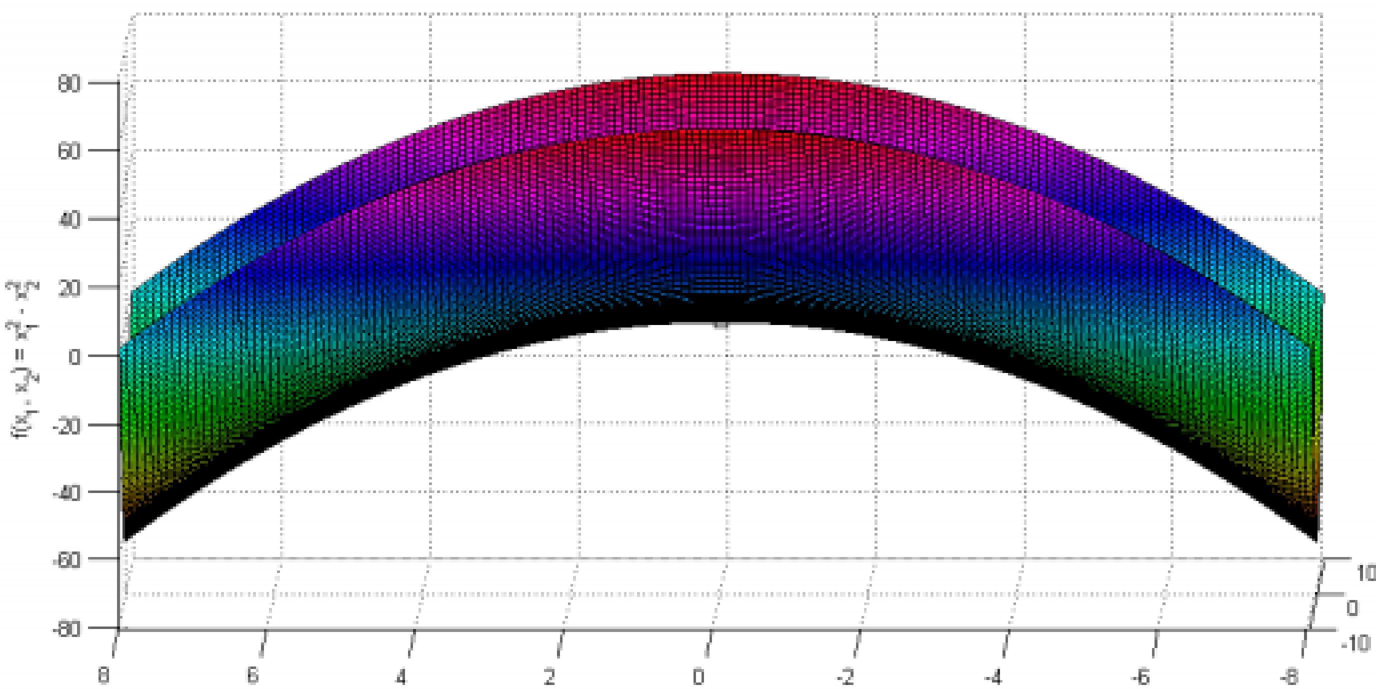

Figure 4.21: The hyperbolic paraboloid $f(x_1, x_2) = x_1^2 - x_2^2$, when vi
the x_2 axis is concave down.

2. (a) Show that a set $S \subseteq \Re^n$ is convex if and only if its intersection with any line is convex.

   (b) Derive a sufficient condition for the matrix $A \in \mathbf{S}^n$ so that the set S specified below (as the solution set of a quadratic inequality) is convex. Assume that $\mathbf{b} \in \Re^n$ and $c \in \Re$.

$$S = \left\{ \mathbf{x} \in \Re^n \,\middle|\, \mathbf{x}^T A \mathbf{x} + \mathbf{b}^T \mathbf{x} + c \leq 0 \right\}$$

   Is this condition on A also necessary? Explain.

(6 Marks)

Soln: (a) Eqn of line is:

$L_{a,b} = \{x = a + \lambda b\}$ where a & b are two pts on the line $(\lambda \in \mathbb{R})$

$\Rightarrow$ Say S is convex. To show $S \cap L_{ab}$ is convex for any a, b

Let $x_1, x_2 \in S \cap L_{ab}$ & $\theta \in [0,1]$

Then $\theta x_1 + (1-\theta) x_2 \in S$

& $\theta x_1 + (1-\theta) x_2 = a + (\lambda_1 \theta + \lambda_2 (1-\theta)) b \in L_{ab}$

$\Rightarrow \theta x_1 + (1-\theta) x_2 \in S \cap L_{ab}$

$\Leftarrow$ Say $S \cap L_{ab}$ is convex $\forall a, b$ & $x_1, x_2 \in S$

Consider $L_{x_1 x_2}$ ... $S \cap L_{x_1 x_2}$ is convex

$\Rightarrow \theta x_1 + (1-\theta) x_2 \in S \cap L_{x_1,x_2} \ \forall \theta \in [0,1]$

$\Rightarrow \theta x_1 + (1-\theta) x_2 \in S \quad \forall \theta \in [0,1]$

$\Rightarrow S$ is convex

(b) We use result in (a):

$S = \{x \in \mathbb{R}^n \mid x^T A x + b^T x + c \leq 0\}$ is convex iff its intersection with any line $L_{u,v}$ is convex

$$(u+\lambda v)^T A (u+\lambda v) + b^T (u+\lambda v) + c = p\lambda^2 + q\lambda + r$$

where: $p = v^T A v \quad q = b^T u + 2u^T A v \quad r = c + b^T u + u^T A u$

Thus the intersection

$\{u+\lambda v \mid p\lambda^2+q\lambda+r \leq 0\}$ is convex

if $p \geq 0$ ... This is true $\forall u$ iff

A is p.s.d

However, the converse does not hold.

Eg: $A=-20$ $b=0$ $c=-10$ then $C=\mathbb{R}$

is convex, although A is NOT p.s.d

3. Let $\mathcal{I}_1 = (V_1, +_1, *_1, <>_1)$ and $\mathcal{I}_2 = (V_2, +_2, *_2, <>_2)$ be two inner product spaces (with respective addition, scalar multiplication and inner product as specified in the 4-tuple) over the scalar field of reals.

   (a) For any $C \subseteq V_1$, define the set $polar(C)$ as follows

   $$polar(C) = \{u \,|u \in V_1,\ < u, v >_1 \leq 1,\ \forall\, v \in C\}$$

   Show that C is a closed convex set containing the origin if and only if

   $$C = polar(polar(C))$$

   Thus, the polar helps provide a dual description of a convex set.
   (6 Marks)
   proposition B.2.2 in Nemirovski. Also http://people.orie.cornell.edu/dpw/orie6300/Lectures/lec07.pdf for polar

Soln:

ⓐ $C \subseteq V$ is a closed convex set

Primal description:

$$C = \text{conv}(\text{convex\_spanning\_set}(C))$$

Dual description:

$$C = \{v \in V \mid \langle v, u\rangle \leq 1 \ \forall u \in \text{Polar}(C)\}$$

$$\text{Polar}(C) = \{u \mid \langle u, v\rangle \leq 1 \ \forall v \in C\}$$

Claim: C is closed & convex

$$\Updownarrow$$

$$C = \text{polar}(\text{polar}(C))$$

Assume $0 \in C$

By defn: $y \in \text{Polar}(C), x \in C \Rightarrow \langle y, x\rangle \leq 1$

$\Rightarrow C \subseteq \text{Polar}(\text{Polar}(C))$

Now let $\exists\ \bar{x} \in \text{Polar}(\text{Polar}(C)) \setminus C$

we need the separation thm which states

Since C is non-empty, convex & closed & $\bar{x} \notin C$ $\bar{x}$ can be "strongly separated" from C. That is, $\exists$ b s.t

$$\langle b, \bar{x} \rangle > \sup_{x \in C} \langle b, x \rangle$$

Now $0 \in C \Rightarrow$ the LHS is positive

Using $a = \lambda b$ for some $\lambda > 0$, we can get $\langle a, \bar{x} \rangle > 1 \geq \sup_{x \in C} \langle a, x \rangle$

This is a contradiction since $1 \geq \sup_{x \in C} \langle a, x \rangle$ implies that $a \in \text{Polar}(C)$

But $\langle a, \bar{x} \rangle > 1$ contradicts that $\bar{x} \in \text{Polar}(\text{Polar}(C))$

(b) Consider the direct sum of the inner product spaces $\mathcal{I}_1 \oplus \mathcal{I}_2$ given by $\mathcal{I}_1 \oplus \mathcal{I}_2 = (V_1 \times V_2, +_3, *_3, <>_3)$ such that, for all $v_{11}, v_{12} \in V_1$ and $v_{21}, v_{22} \in V_2$ and $\alpha \in \Re$

$$\alpha *_3 (v_{11}, v_{21}) = (\alpha *_1 v_{11}, \alpha *_2 v_{21})$$

and

$$(v_{11}, v_{21}) +_3 (v_{12}, v_{22}) = (v_{11} +_1 v_{12}, v_{21} +_2 v_{22})$$

Let $K_1 \subset V_1$ and $K_2 \subset V_2$ be closed convex cones. Now answer the following questions:

i. Is $K_1 \times K_2$ a closed convex cone?

ii. Write an expression for the dual cone $(K_1 \times K_2)^*$ in terms of the dual cones K_1^* and K_2^*. Prove that your answer is correct.

(7 Marks)

Soln: (1) Yes. For any θ & $\lambda \geq 0$, $V_{11}, V_{12} \in K_1$, $V_{21}, V_{22} \in K_2$

$$\theta \cdot_3 (V_{11}, V_{21}) +_3 \lambda \cdot_3 (V_{12}, V_{22}) = (\underbrace{\theta \cdot_1 V_{11} +_1 \lambda \cdot_1 V_{12}}_{\in K_1}, \underbrace{\theta \cdot_2 V_{21} +_2 \lambda \cdot_2 V_{22}}_{\in K_2})$$

$$\in K_1 \times K_2$$

$\Rightarrow$ $K_1 \times K_2$ is a convex cone. It is also closed

(2) $(K_1 \oplus K_2)^* = (K_1^* \oplus K_2^*)$

We prove that

$x \in (K_1 \oplus K_2)^*$ iff $\langle x, v \rangle \geq 0 \;\forall\; v \in K_1 \times K_2$

$(x = (x_1, x_2), v = (v_1, v_2))$ ie iff $\langle x_1, v_1 \rangle_1 + \langle x_2, v_2 \rangle_2 \geq 0 \;\forall\; v_1 \in K$ & $v_2 \in K_2$ (A)

ie iff $x_1 \in K_1^*$ & $x_2 \in K_2^*$. (B)

That B $\Rightarrow$ A is obvious since

$x_1 \in K_1^*$ & $x_2 \in K_2^* \Rightarrow \langle x_1, v_1\rangle_1 \geq 0$ & $\langle x_2, v_2\rangle_2 \geq 0$
$\forall\ v_1 \in K_1$ & $v_2 \in K_2$

$$\Rightarrow \langle x_1, v_1\rangle_1 + \langle x_2, v_2\rangle_2 = \langle x, v\rangle_3 \geq 0$$

$$\Rightarrow x \in (K_1 \oplus K_2)^*$$

To prove A $\Rightarrow$ B, ie $\langle x_1, v_1\rangle_1 + \langle x_2, v_2\rangle_2 \geq 0 \ \forall\ v_1 \in K_1$, $v_2 \in K_2$

$$\Rightarrow \langle x_1, v_1\rangle_1 \geq 0 \quad \& \quad \langle x_2, v_2\rangle_2 \geq 0$$

Taking $v_1 = 0$ (since $0 \in K_1$), we get from A that

$$\langle x_1, 0\rangle_1 + \langle x_2, v_2\rangle_2 \geq 0 \quad \forall\ v_2 \in K_2$$

$$\Rightarrow \langle x_2, v_2\rangle_2 \geq 0 \quad \forall\ v_2 \in K_2$$

Taking $v_2 = 0$ (since $0 \in K_2$) we get from B that

$$\langle x_1, v_1\rangle_1 + \langle x_2, 0\rangle_2 \geq 0 \quad \forall\ v_1 \in K_1$$

$$\Rightarrow \langle x_1, v_1\rangle_1 \geq 0 \quad \forall\ v_1 \in K_1$$

Hence proved iff condition

4. Express the following problem as a conic program and write its dual as compactly as possible: Find the minimum eucledian distance to a given affine subspace of $\Re^n$ from the origin ($\mathbf{0} \in \Re^n$).

(5 Marks)

Soln:

$$\min_{\substack{x \in \mathbb{R}^n \\ Ax = b}} \|0 - x\|_2 \equiv \min_{\substack{\|x\|_2 \le r \\ Ax = b \\ x \in \mathbb{R}^n}} r \equiv \min_{\substack{y \in K \\ \begin{bmatrix} A & 0 \end{bmatrix}\begin{bmatrix} x \\ r \end{bmatrix} = \begin{bmatrix} b \\ 0 \end{bmatrix}}} c^T y$$

$y = \begin{bmatrix} x \\ r \end{bmatrix}$ $c = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$

strictly speaking, not a cone (linear cone) program

$\{(x, r) \mid \|x\|_2 \le r\} = K = L\text{-}2$ norm cone

$\{(x, r) \mid \|x\|_2 \le r\} = K^* = L2$ norm cone

Dual: $\max_{\lambda} b^T \lambda$ $\|A^T \lambda\|_2 \le 1$

5. In the class, we gave an analytic proof for the strong duality theorem for Conic Programs. In this question, we will attempt to give a geometrically motivated proof (under certain assumptions) for the special case of linear programs 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)

$$\begin{array}{ll} \min\limits_{\mathbf{x} \in \Re^n} & \mathbf{c}^T \mathbf{x} \\ \text{subject to} & A\mathbf{x} \geq \mathbf{b} \end{array} \tag{1}$$

   and D be the dual program given in (2)

$$\begin{array}{ll} \max\limits_{\mathbf{y} \in \Re^m} & \mathbf{b}^T \mathbf{y} \\ \text{subject to} & A^T \mathbf{y} = \mathbf{c} \\ & \mathbf{y} \geq \mathbf{0} \end{array} \tag{2}$$

   Assume that the dual D is feasible. Also, 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 $K = \left\{ \mathbf{x} \;\middle|\; \mathbf{x} = \sum\limits_{i \in I} \lambda_i \mathbf{a}_i, \lambda_i \geq 0 \right\}$ generated by the set of vectors $\{\mathbf{a}_i\}_{i \in I}$. 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 $\epsilon \leq \min\limits_j \left(\frac{b_j - \mathbf{a}_j^T \mathbf{x}^*}{\mathbf{a}_j^T \mathbf{x}^*} \right)$, we have $\mathbf{a}_j^T \mathbf{z} = \mathbf{a}_j^T \mathbf{x}^* + \epsilon \mathbf{a}_j^T \mathbf{d} \geq b_j$.

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

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)