Midsem 2015

37 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. Unless I ask you to prove something stated in class, you can assume facts proved or stated in class without proof.

1. Find and classify (as local or global maximum or minimum or as a saddle point) the stationary points for the following function.

$$f(x) = (x_2 - x_1^2)^2 + x_1^5 = x_2^2 + x_1^4 + x_1^5 - 2x_2x_1^2 = x_2(x_2 - 2x_1^2) + x_1^4 + x_1^5$$

(4 Marks)

$$\nabla f(x) = \begin{bmatrix} 5x_1^4 + 4x_1^3 - 4x_2x_1 \\ 2x_2 - 2x_1^2 \end{bmatrix} = \begin{bmatrix} 5x_1^4 + 4x_1(x_1^2 - x_2) \\ 2(x_2 - x_1^2) \end{bmatrix}$$

Solving for $\nabla f(x^*) = 0$ we get

$x_2 = x_1^2$ & $x_1^4 = 0 \Rightarrow x^* = 0$ as only critical point

$$\nabla^2 f\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right) = \begin{bmatrix} 0 & 0 \\ 0 & 2 \end{bmatrix}$$ which is p.s.d

Note from theorem 61 of http://www.cse.iitb.ac.in/~cs709/notes/BasicsOfConvexOptimization.pdf

that a sufficient condition for local min is that Hessian is positive definite (whereas $\nabla^2 f\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}\right)$ is only p.s.d)

Thus, we need to find other ways of determining

If $\begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is local min or max

Consider some pt (λ_1, λ_2) in the neighborhood of $(0,0)$

$$f(\lambda_1, \lambda_2) = \lambda_2(\lambda_2 - 2\lambda_1^2) + \lambda_1^4 + \lambda_1^5$$

Taking cue from the form of f, consider values along the curve $\lambda_2 = \lambda_1^2 + \lambda_1^3$

ie $f(\lambda_1, \lambda_1^2 + \lambda_1^3) = \lambda_1^6 - \lambda_1^4 + \lambda_1^4 + \lambda_1^5 = \lambda_1^6 + \lambda_1^5$

We note that for small λ_1, $f(\lambda_1, \lambda_1^2 + \lambda_1^3) < 0$ if $\lambda_1 < 0$

$f(\lambda_1, \lambda_1^2 + \lambda_1^3) > 0$ if $\lambda_1 > 0$

& $f(\lambda_1, \lambda_1^2 + \lambda_1^3) = 0$ if $\lambda_1 = 0$

$\Rightarrow x^* = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ is a saddle pt

2. Find $\beta, \theta \in \Re$ for which the function $f(x, y) = \beta(x^2+y^2)+\theta xy+x+y$
   (a) has no stationary points
   (1 Mark)
   (b) has exactly one stationary point and it is a global strict minimum
   (1 Mark)
   (c) has infinite stationary points, and all of them are global minimizers
   (1 Mark)

Ans: $\nabla f(x,y) = \begin{bmatrix} 2\beta x+\theta y+1 \\ 2\beta y+\theta x+1 \end{bmatrix} = 0$ requires

$2\beta x+\theta y = 2\beta y+\theta x = -1$

if $\beta=\theta=0$ then this is impossible

(a) No stationary pt

Else: if $2\beta+\theta \neq 0$ & $2\beta-\theta \neq 0$ then $x=y=\dfrac{-1}{2\beta+\theta}$ is exactly one stationary pt

Further, $\nabla^2 f(x,y) = \begin{bmatrix} 2\beta & \theta \\ \theta & 2\beta \end{bmatrix}$

$\lambda_1+\lambda_2 = 4\beta$

$\lambda_1 * \lambda_2 = 4\beta^2-\theta^2$

$\therefore$ if $4\beta>0$ & $4\beta^2-\theta^2>0$ ie $\beta^2 > \dfrac{\theta^2}{4}$

then f is strictly convex & $\begin{bmatrix} \dfrac{-1}{2\beta+\theta} \\ \dfrac{-1}{2\beta+\theta} \end{bmatrix}$ is global minimum (strict) (b)

Else if $\theta = 2\beta$ then $\theta(x+y) = -1$ corresponds to infinite stationary pts. Further, if (like before) $\beta^2 > \theta^2/4$ & $4\beta>0$ then we have infinite global minima (c)

3. Consider the quadratic function

$$f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} + \mathbf{x}^T b + c \qquad (1)$$

(a) Suppose you are told that the quadratic function is convex in a particular convex domain $\mathcal{D}$. That is, $f(\mathbf{x})$ is convex when $\mathbf{x} \in \mathcal{D}$. Is it necessary that $f(\mathbf{x})$ be convex in any other convex domain $\mathcal{E}$?
What about strict and strong convexity?

(b) What if $f(\mathbf{x})$ were any arbitrary function, convex in a convex domain $\mathcal{D}$? Is it necessary that $f(\mathbf{x})$ is also convex in any other convex domain $\mathcal{E}$?
Prove your claims.

(**3 Marks**)

i.e. not lie in an affine subspace of dim < n

Ans: Strictly speaking, the answer below requires even in quad case that D should NOT have an empty interior in $\mathbb{R}^n$ with dimension less than n [Any one who points this out gets upto 2 more bonus marks]

Eg: if $D = \{[x_1\ 0 \dots 0] \mid x_1 \in \mathbb{R}\}$ $f(x) = a_{11}x_1^2 + b_1 x_1 + c$ which is (strictly) convex in x_1 if $a_{11} > 0$

(b) However $\nabla^2 f(x) = A$ may not be positive definite & ∴ $f(x)$ may not be strictly convex on $\mathbb{R}^n$ & it's several subsets even if $f(x)$ is convex on $D = \{(x_1, \dots 0] \mid x_1 \in \mathbb{R}\}$

Now suppose D has non-empty interior in $\mathbb{R}^n$. Then

(a) f is differentiable & doubly differentiable in the interior of D & $\therefore$ $\nabla^2 f(x) = A$

if f is convex on D then $\nabla^2 f(x) = A \succeq 0$

$\Rightarrow$ f is convex on any subset of $\mathbb{R}^n$

Ditto for strict convexity & strong convexity

4. The set of subgradients of a function f at a point $\mathbf{x}$ is called its subdifferential $\partial f(\mathbf{x})$.

   (a) Prove that the subdifferential $\partial f(\mathbf{x})$ is a closed convex set.

   (b) Prove that if $f(\mathbf{x}) = \alpha_1 f_1(\mathbf{x}) + \alpha_2 f_2(\mathbf{x})$ with $\alpha_1, \alpha_2 \geq 0$, then

   $$\partial f(\mathbf{x}) = \alpha_1 \partial f_1(\mathbf{x}) + \alpha_2 \partial f_2(\mathbf{x})$$

   where RHS is element-wise addition between two sets. That is,

   $$\alpha_1 \partial f_1(\mathbf{x}) + \alpha_2 \partial f_2(\mathbf{x}) = \{\mathbf{g} \mid \mathbf{g} = \alpha_1 \mathbf{g}_1 + \alpha_2 \mathbf{g}_2,\ \mathbf{g}_1 \in \partial f_1(\mathbf{x}),\ \mathbf{g}_2 \in \partial f_2(\mathbf{x})\}$$

Ans: (a) $\partial f(x) = \{ g \mid f(y) \geq f(x) + g^T(y-x) \quad \forall\, y \in \mathrm{dmn} f \}$

if $g_1, g_2 \in \partial f(x)$ & $\alpha \in [0,1]$

then $f(x) + (\alpha g_1 + (1-\alpha) g_2)^T (y-x)$

$= \alpha\left(f(x) + g_1^T(y-x)\right) + (1-\alpha)\left(f(x) + g_2^T(y-x)\right)$

$\leq \alpha f(y) + (1-\alpha) f(y) \quad [\because g_1, g_2 \in \partial f(x)]$

$= f(y)$

$\Rightarrow \alpha g_1 + (1-\alpha) g_2 \in \partial f(x)$ ie $\partial f(x)$ is convex

Further $\partial f(x)$ is intersection of infinite half spaces, each being a closed (convex) set .. ie

$$\partial f(x) = \bigcap_{y \in \mathrm{dmn} f} \{ g \mid f(y) \geq f(x) + g^T(y-x) \}$$

$\Rightarrow \partial f(x)$ is closed

(b) Proof in two parts (the latter is non-trivial & involves separation theorem & therefore will be given 3 extra marks if solved. Further, the latter requires more assumptions. So a counter example to disprove will fetch 2 extra marks)

(i) $\alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x) \subseteq \partial(\alpha_1 f_1 + \alpha_2 f_2)(x)$

If $g_1 \in \partial f_1(x)$ & $g_2 \in \partial f_2(x)$ then $\forall y \in \operatorname{dmn} f_1 \cap \operatorname{dmn} f_2$

$$f_1(y) \geq f_1(x) + g_1^t(y-x) \quad * \ \alpha_1 \geq 0$$

$$+ \quad f_2(y) \geq f_2(x) + g_2^t(y-x) \quad * \ \alpha_2 \geq 0$$

$$\alpha_1 f_1(y) + \alpha_2 f_2(y) \geq (\alpha_1 f_1(x) + \alpha_2 f_2(x)) + (\alpha_1 g_1^t + \alpha_2 g_2^t)(y-x)$$

$$\Rightarrow \alpha_1 g_1 + \alpha_2 g_2 \in \partial(\alpha_1 f_1 + \alpha_2 f_2)(x)$$

(ii) $\partial(\alpha_1 f_1 + \alpha_2 f_2)(x) \subseteq \alpha_1 \partial f_1(x) + \alpha_2 \partial f_2(x)$

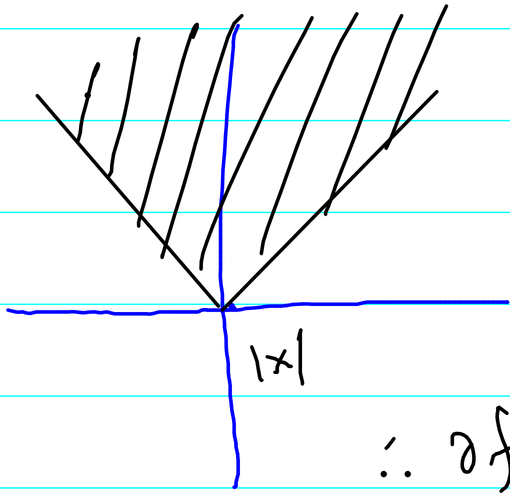
Let $g \in \partial(\alpha_1 f_1 + \alpha_2 f_2)(x)$

Let $S_{f_1} = \left\{(y-x, z) \in \mathbb{R}^n \times \mathbb{R} \mid z > f_1(y) - f_1(x) - g^T(y-x)\right\}$

$S_{f_2} = \left\{(y-x, z) \in \mathbb{R}^n \times \mathbb{R} \mid -z > f_2(y) - f_2(x) - g^T(y-x)\right\}$

Requires $\operatorname{int}(\operatorname{dmn} f_1 \cap \operatorname{dmn} f_2) \neq \phi$

(c) Derive a generic expression for any element of $\partial f(0)$ for $f(x) = |x|$ where $x \in \Re$.



$\partial f(0)$ is set of all supporting hyperplanes to shaded epigraph at 0

$$\therefore\ \partial f(0) = \left\{ (x_1, x_2) \,\middle|\, x_2 = \alpha x_1 \quad \alpha \in \left[-\tfrac{1}{2}, \tfrac{1}{2}\right] \right\}$$

(d) For a non-differentiable (but let us say continuous on its convex domain D) convex function $f(\mathbf{x})$ defined on $\mathbf{x} \in D$, derive and prove a neccessary and sufficient condition for global minimum at $\mathbf{x}^* \in D$ in terms of some condition on $\partial f(\mathbf{x}^*)$.

$\forall\, g \in \partial f(x^*),\quad f(y) \geq f(x^*) + g^T(y - x^*)$ (1)

Needed: $f(y) \geq f(x^*)\ \forall\ y \in D$ (2)

Obviously: $g = 0 \in D$ will be a sufficient condition for (2) given (1)

From (1), in fact $\boxed{g^T(y - x^*) \geq 0\ \forall\, y \in D}$ is itself a sufficient condition

Next: Suppose $f(y) \geq f(x^*)\ \forall\ y \in D$ ie (2) holds

Then obviously (1) must hold for $g = 0$. Thus $g = 0$ ie $0 \in \partial f(x^*)$ is a necessary condition as well for x^* to be a pt of global minimum

$\therefore$ $0 \in \partial f(x^*)$ is a necessary & sufficient condition

(e) Prove the conjugate subgradient theorem, which states that if f is a closed convex function, then for any $\mathbf{x}_1, \mathbf{x}_2$

$$\langle \mathbf{x}_1, \mathbf{x}_2 \rangle = f(\mathbf{x}_1) + f^*(\mathbf{x}_2)$$

holds if and only if

$$\mathbf{x}_2 \in \partial f(\mathbf{x}_1)$$

where $f^*(\mathbf{x})$ is the convex conjugate of $f(\mathbf{x})$

$$f^*(x_2) = \sup_{x \in D} \{ \langle x, x_2 \rangle - f(x) \}$$

1 $\rightarrow f^*(x_2) \geq \langle x, x_2 \rangle - f(x) \ \forall x$

We have stated in class that if f is closed convex (i.e epi f is closed) then

$$f(x_1) = \sup_{x} \{ \langle x_1, x \rangle - f^*(x) \}$$

2 $\rightarrow f(x_1) \geq \langle x_1, x \rangle - f^*(x) \ \forall x$

From 1 & 2 $x_2 \in \partial f(x_1) \Longleftrightarrow x_1 \in \partial f^*(x_2)$

i.e $f(x_1) + f^*(x_2) = \langle x_1, x_2 \rangle$ iff $x_2 \in \partial f(x_1)$

(10 Marks)

5. Consider a constrained convex optimization problem:

$$\begin{aligned} &\text{minimize} && f(\mathbf{x}) \\ &\text{subject to} && g_i(\mathbf{x}) \leq 0 \quad \text{for } i = 1 \ldots m \end{aligned} \tag{2}$$

where f and g_i's are closed convex functions.

We will discuss two ways of reformulating this problem:

(a) ~~Consider indicator function $I_{g_i}(\mathbf{x})$ associated with $g_i(\mathbf{x})$ for each $i =$~~ $1 \ldots m$ such that $I_{g_i}(\mathbf{x}) = 0$ iff $g_i(\mathbf{x}) \leq 0$ and $I_{g_i}(\mathbf{x}) = 1$ otherwise.

i. Prove that $\partial I_{g_i}(\mathbf{x})$ is a convex cone. Is it closed?

ii. Pose (2) as an equivalent unconstrained convex optimization problem making use of $I_{g_i}(\mathbf{x})$.

iii. ~~Now derive a necessary and sufficient condition for global con-~~strained optimality of (2) at a point $\mathbf{x}^*$.

(5 Marks)

Let $S_{g_i} = \{y \mid g_i(y) \leq 0\}$

S_{g_i} is a convex fn (sublevel set of convex fn)

(a) (i) We can prove that $\partial I_{g_i}(x) = \{g \mid g^T y \leq g^T x \;\; \forall y\}$

$I_{g_i}(y) \geq I_{g_i}(x) + \text{subgrad}^T(x)(y-x) \quad \forall y$

If: $g_i(x) \leq 0$ then $0 \geq \text{subgrad}^T(x)(y-x)$ is necessary & sufficient

ie $\partial I_{g_i}(x) = \{g \mid g^T y \leq g^T x \;\; \forall y \in S_{g_i}\} = \bigcap_{y \in S_{g_i}} \{g \mid g^T(y-x) \leq 0\}$ = intersection of half spaces of infinite hyperplane through origin = closed convex cone

$\therefore I_{g_i}(x)$ is a closed convex cone $\forall \; x : g_i(x) \leq 0$

$I_{g_i}(x) = \phi$ if $x : g_i(x) > 0$

(ii) There are multiple ways to achieve it:-

$\lim_{\lambda \to \infty} \min_x f(x) + \lambda \sum_i I_{g_i}(x)$ ... Equivalent to redefining

$I_{g_i}(x) = 0$ if $g_i(x) \leq 0$

$= \infty$ o/w

& using $\min_x f(x) + \sum_i I_{g_i}(x)$

OR

$\min_x \max_{\lambda_i} f(x) + \sum_i \lambda_i I_{g_i}(x)$

(iii) Necessary & sufficient conditions for optimality

(from Q4) : Let $I_{g_i}(x) = 0$ if $g_i(x) \leq 0$

$= \infty$ o/w

$0 \in \partial\left(f(x) + \sum I_{g_i}(x)\right)$

$\partial I_{g_i}(x) = \{g \mid g^T y \leq g^T x \;\; \forall y : g_i(y) \leq 0\}$ if $g_i(x) \leq 0$

& $\partial I_{g_i}(x) = \phi$ o/w

(b) We define the following Conic Linear Program (Conic LP):

$$\begin{aligned} \text{minimize} \quad & \mathbf{c}^T\mathbf{x} \\ \text{subject to} \quad & \mathbf{x} \in \mathcal{K} \\ & A\mathbf{x} = \mathbf{b} \end{aligned} \tag{3}$$

where $\mathcal{K}$ is a closed convex cone. Prove that corresponding to any convex optimization problem of the form (2), there exists a corresponding Conic LP (3) which has the same solution. Your proof is by deriving the conic program corresponding to the convex optimization problem.

Extra and Optional: Now how would you derive the convex conjugate for the Conic LP (3). What is its relation with the dual cone $\mathcal{K}^*$?

(**6 Marks**)

Claim: $C = \{x \mid g_i(x) \le 0 \ \forall x\}$

$K = \{(x, y) \mid y > 0 \ \ \frac{x}{y} \in C\}$ Is $\underbrace{\text{a cone}}_{}$

Need to add $\{0, 0\}$

$$\frac{\alpha_1 x_1 + \alpha_2 x_2}{\alpha_1 y_1 + \alpha_2 y_2} = \frac{\alpha_1 y_1}{\alpha_1 y_1 + \alpha_2 y_2}\left(\frac{x_1}{y_1}\right) + \frac{\alpha_2 y_2}{\alpha_1 y_1 + \alpha_2 y_2}\left(\frac{x_2}{y_2}\right) \in C$$

We will have:

$$g_i(x) \le 0 \ \ \forall i \iff (x, 1) \in K = \{(x, y) \mid y > 0 \ \ g_i(x/y) \le 0 \ \forall i\}$$

Given

$$\min_x f(x)$$

$$\text{s.t } g_i(x) \le 0$$

We equivalently formulate

$$\min_{x,t} \; t$$

$$f(x) - t \leq 0$$

$$g_i(x) \leq 0$$

Or equivalently

$$\min_{x,t} \; t$$

$$(x, t, 1) \in \underbrace{\left\{ (x, t, y) \;\middle|\; \begin{array}{l} f(x/y) - t/y \leq 0 \\ g_i(x/y) \leq 0 \quad \forall i \\ y > 0 \end{array} \right\} \cup \{(0, 0, 0)\}}_{\text{A convex cone } K}$$

10

(6) 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 $\langle x_1, v_1\rangle_1 + \langle x_2, v_2\rangle_2 \geq 0 \ \forall\ v_1 \in K_1$, $v_2 \in K_2$

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

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