1. Let $f : \Re^n \rightarrow \Re$ be a convex function and let $\mathcal{C} \subseteq \Re^n$ be such that the following set

$$C_\delta = \{\mathbf{z} \mid ||\mathbf{z} - \mathbf{x}|| \leq \delta, \ \textit{for some } \mathbf{x} \in \textit{closure}(\mathcal{C})\}$$

is compact[1], that is, C_δ is closed and bounded, for all $\delta \geq 0$. Show that f is Lipschitz continuous on $\mathcal{C}$, i.e., there exists $L > 0$ such that

$$|f(\mathbf{x}) - f(\mathbf{y})| \leq L||\mathbf{x} - \mathbf{y}||, \ \forall x, y \in \mathcal{C}$$

State the value for L.

Hint: Use the Weierstrass' theorem, which implies that a continuous function $f : \Re^n \rightarrow \Re$ over a compact set $\mathcal{A} \subseteq \Re^n$ attains extreme values on $\mathcal{A}$.

(7 Marks)

Ans: (1) As discussed in class, a convex fn on $\mathbb{R}^n$ is continuous. Thus, f is continuous

(2) Let $p, q \in C_\delta$ then

$$r = q + \frac{\delta}{||q-p||}(q-p) \in C_\delta$$ (By definition of C_δ)

$$\Rightarrow q = \underbrace{\frac{||q-p||}{||q-p||+\delta}}_{\alpha} r + \underbrace{\frac{\delta}{||q-p||+\delta}}_{1-\alpha} p$$

Thus, q is a convex combination of p & r

Since f is convex:

$$f(q) \leq \underbrace{\frac{||q-p||}{||q-p||+\delta}}_{\alpha} f(r) + \underbrace{\frac{\delta}{||q-p||+\delta}}_{1-\alpha} f(p)$$

$$\Rightarrow f(q) - f(p) \le \frac{\|q-p\|}{\|q-p\|+\delta}\big(f(r) - f(p)\big)$$

$$\frac{\|q-p\|}{\|q-p\|+\delta} \le \frac{\|q-p\|}{\delta}$$

[By Weierstrass' thm, f is cts & since C_δ is closed & bounded, f attains extrema on C_δ]

$$\le \frac{\|q-p\|}{\delta}\left(\max_{z\in C_\delta} f(z) - \min_{x\in C_\delta} f(x)\right) \quad (1)$$

Similarly, we get

$$f(p) - f(q) \le \frac{\|p-q\|}{\delta}\left(\max_{z\in C_\delta} f(z) - \min_{x\in C_\delta} f(x)\right) \quad (2)$$

Combining (1) & (2), we get

$$|f(p) - f(q)| \le L\|p-q\|$$

where $L = \left(\max_{z\in C_\delta} f(z) - \min_{x\in C_\delta} f(x)\right)\Big/\delta$

2. As discussed in the class, directon $\mathbf{d}$ is a descent direction of a function f at a point $\mathbf{x}$ if the directional derivative of f along $\mathbf{d}$ is strictly negative. That is $\mathbf{d}^T\nabla f(\mathbf{x}) < 0$. In this exercise, we provide a method for generating descent direction in cases in which obtaining a single subgradient is relatively simple.

(a) Let $\mathbf{g}_f^{(i)}(\mathbf{x})$ be a subgradient of f at $\mathbf{x}$ in the i^{th} step of the algorithm. (For $i = 0$, you just pick any subgradient.) Let $\mathbf{w}_k$ be the vector of minimum p-norm (for any $p \geq 1$) in the convex hull of $\mathbf{g}_f^{(1)}(\mathbf{x}), \mathbf{g}_f^{(2)}(\mathbf{x}), \ldots \mathbf{g}_f^{(k-1)}(\mathbf{x})$. Present an algorithm for computing $\mathbf{w}_k$ when $p = 2$. What about the case of any other value of p?

$\ldots \mathbf{g}_f(\mathbf{x}^{(k-1)})$.

(4 Marks)

Ans: $W_k = \min_{\alpha, w} \|W\|_p$

s.t. $\underbrace{w = \alpha_1 g_f^{(1)}(x) + \alpha_2 g_f^{(2)}(x) \cdots + \alpha_{k-1} g_f^{(k-1)}(x)}$

$$[W^T \; -\alpha_1 \; -\alpha_2 \; \cdots \; -\alpha_{k-1}] \begin{bmatrix} I_{n\times n} \\ g_f^{(1)}(x) \\ g_f^{(2)}(x) \\ \vdots \\ g_f^{(k-1)}(x) \end{bmatrix} = 0$$

Matrix M

ie $\underbrace{[W^T \; -\alpha_1 \; -\alpha_2 \; \cdots \; -\alpha_{k-1}] \in \text{Nullspace}(M) = N_{basis} Y}$

Substitute to get an convex unconstrained optimization problem & solve using gradient descent etc

Holds for $p=2$ or even otherwise

(b) Stop if $-\mathbf{w}_k$ is a descent direction of f at $\mathbf{x}$. Since f may not be differentiable, the criterion for $-\mathbf{w}_k$ being a descent direction of f at $\mathbf{x}$ is obtained by replacing $\nabla f(\mathbf{x})$ with $\mathbf{g}_f(\mathbf{x})$:

$$-\mathbf{w}_k^T \mathbf{g}_f(\mathbf{x}) < 0$$

If the stopping criterion is not met, let $\mathbf{g}_f^{(k)}(\mathbf{x}) \in \partial f$ such that

$$w_k^T \mathbf{g}_f^{(k)}(\mathbf{x}) = \min_{\mathbf{g} \in \partial f} w_k^T \mathbf{g}$$

Prove that this process returns a descent direction of f at $\mathbf{x}$ in a finite number of iterations. You can assume that ∂f is compact. (Note that since $\partial f \subseteq \Re^n$, this is equivalent to saying that ∂f is closed and bounded).

(7 Marks)

Ans: Firstly, w_k is the projection of the origin on the set $\text{conv}(g_f^{(1)}(x), g_f^{(2)}(x), \ldots g_f^{(k-1)}(x))$

$\therefore$ By the projection theorem on slide 9 of

http://www.cse.iitb.ac.in/~CS709/notes/eNotes/first-order-descent-projectionMethod-annotated.pdf, we have $\forall\, g \in \text{conv}(g_f^{(1)}(x) \ldots g_f^{(k-1)}(x))$

$$(g - w_k)^T (0 - w_k) \leq 0$$

$$\Rightarrow (g - w_k)^T w_k \geq 0 \Rightarrow g^T w_k \geq \|w_k\|^2 \geq \underbrace{\min_{g \in \partial f(x)} \|g\|^2}_{} = \|g_*\|^2$$

Note: $w_k \in \partial f(x)$

& $g_* \in \partial f(x)$ is subgradient with minimum norm

Both $(g_f^{(1)} \ldots g_f^{(k-1)})$ & $(w_1 \ldots w_{k-1})$ are sequences lying in ∂f

- If $\|g_x\| = 0$ then x should be a minimizer of f & we are already at optimal soln

Else $\|g_x\| > 0 \Rightarrow \forall g \in \partial f(x)$, $g^T w_k > 0$

- ∂f is closed $\Rightarrow$ The sequences $\{w_1 \ldots w_k \ldots\}$ & $\{g_f^{(1)} \ldots g_f^{(k)} \ldots\}$ must have limit points $\hat{w}$ & $\hat{g}_f$ in ∂f

$\Rightarrow \hat{g}^T \hat{w} > 0 \longrightarrow$ 1

However, since none of $w_1 \ldots w_k$ have been descent directions, $g_f^{(1)T}(-w_1) \geq 0$ & $\ldots$ $g_f^{(k-1)T}(-w_{k-1}) \geq 0$

and as $k \to \infty$ $\quad \hat{g}_f^T(-\hat{w}) \geq 0 \longrightarrow$ 2

1 and 2 directly contradict each other

$\Rightarrow$ Process MUST terminate with a descent direction in a finite number of steps

Proof again summarized on next slide

$\Rightarrow$ Since ∂f is closed, limit points of $w_1 \dots w_{k-1}$ & of $(g_f^{(1)} \dots g_f^{(k-1)})$ should both lie in ∂f

$\|g_x\| > 0$, since if $g_x = 0$ then x is already an optimal point

$\therefore$ if $\lim_{k\to\infty} w_k = \hat{w}$ & $\lim_{k\to\infty} g_f^{(k)} = \hat{g}_f$ then we expect $\hat{g}_f^T \hat{w} > 0$

But if none of w_k's are descent directions so far, we have

$$-w_k^T g_f^{(k)} = -\min_{g\in\partial f} w_k^T g = \max_{g\in\partial f} -w_k^T g \geq 0$$

ie $w_k^T g_f^{(k)} \leq 0$

$\hookrightarrow$ Letting $k\to\infty$, we get a contradiction

$$\hat{w}^T \hat{g}_f \leq 0$$

$\Rightarrow$ Some w_k should become a descent direction

3. Consider the equality constrained optimization problem in (1):

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{x}^T Q\mathbf{x} + \mathbf{c}^T\mathbf{x} + \beta \\ \text{subject to} & A\mathbf{x} = \mathbf{b} \end{array} \tag{1}$$

Assume that A has full row rank (that is no equality is redundant/conflicting). Let N be the basis for the null space of A. Show that this optimization problem is unbounded below if $N^T QN$ has negative eigenvalues.

(5 Marks)

Ans: Let us write KKT conditions for this problem

$$Qx^* + c + A^T\lambda^* = 0$$
$$Ax^* = b$$

Let N^TQN have a negative eigenvalue λ with corresponding eigenvector v ... ie $N^TQNv = \lambda v$, $\lambda \leq 0$

Then $\overbrace{\underbrace{v^TN^T}_{u^T} Q \underbrace{Nv}_{u}}^{a} < 0$ & $u = Nv \Rightarrow Au = 0$

$\Rightarrow \forall \alpha \neq 0$ $A(x^* + \alpha u) = b$ .. ie $x^* + \alpha u$ is feasible

However: The value of the objective at $x^* + \alpha u$ is

$f(x)$

$$f(x^* + \alpha u) = f(x^*) + \alpha u^T(Qx^* + c) + \frac{1}{2}\alpha^2 u^TQu$$

$$= f(x^*) - \underbrace{\alpha u^T A^T \lambda^*} + \frac{1}{2}\alpha^2 u^T Q u$$

Since as per KKT condition,

$$Qx^* + C = -A^T\lambda^*$$

$$= f(x^*) + \frac{1}{2}\alpha^2 u^T Q u \quad \left(\because u^T A^T \lambda^* = v^T N^T A^T \lambda^* = 0 \text{ by defn of null space} \right)$$

$$< f(x^*) \quad \left(\because u^T Q u < 0 \text{ from } ⓐ \right)$$

$\therefore$ For any x^* satisfying the KKT condition, we can find a feasible direction along u along which f strictly decreases

Now consider the inequality constrained optimization problem in (2) and the primal active set method for the same that we had discussed in the class. If the same unboundedness problem persists in this case, then the algorithm might never terminate. Can that actually happen if A has full row rank and N^TQN has negative eigenvalues? Explain.

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}\mathbf{x}^TQ\mathbf{x} + \mathbf{c}^T\mathbf{x} + \beta \\ \text{subject to} & A\mathbf{x} \geq \mathbf{b} \end{array} \tag{2}$$

Soln: (... please use your own argument)

With reference to step (2) of Algorithm in
http://www.cse.iitb.ac.in/~cs709/notes/quadraticOpt-PrimalActiveSet.pdf

It should be possible to show that if A has full row rank than $A_{\mathcal{I}_k}$ (A with rows restricted to active index set $\mathcal{I}_k$) should also have full row rank & if N^TQN has negative eigenvalues then the optimization problem in step (2) of active set algo will be unbounded below

4. Motivate and explain the dual ascent method, the augmented lagrangian method and the alternating direction method of multipliers (ADMM) methods.

(2 Marks)

Refer to pages 9 & 10 of class notes at

http://www.cse.iitb.ac.in/~CS709/notes/eNotes/2015-22.pdf

Recall that the augmented lagrangian was to make the optimization problem strongly convex[2], and therefore improve convergence. We will similarly try to modify the barrier (interior-point) method that was discussed in class.

The general inequality constrained convex minimization problem is

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) \le \mathbf{0}, \quad i = 1, \dots, m \\ & A\mathbf{x} = b \end{array} \tag{3}$$

The Barrier method solves (3) by making a sequence of approximations in terms of solutions to problem (4):

$$\begin{array}{ll} \text{minimize} & B(\mathbf{x}, \mu) = f(\mathbf{x}) - \mu \sum_{i=1}^{m} \ln\left(-g_i(\mathbf{x})\right) \\ \text{subject to} & A\mathbf{x} = b \end{array} \tag{4}$$

The objective function $B(\mathbf{x}, \mu)$ is called the *logarithmic barrier function*. This function is convex, which can be proved by invoking the composition rules.

Now we add the constraint $||\mathbf{x}||^2 \le \rho^2$ to the problem in (3) to get (5)

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & g_i(\mathbf{x}) \le \mathbf{0}, \quad i = 1, \dots, m \\ & A\mathbf{x} = b \\ & ||\mathbf{x}||^2 \le \rho^2 \end{array} \tag{5}$$

Let $\overline{B}(\mathbf{x}, \mu)$ denote the *(modified) logarithmic barrier function* to this modified problem. Prove that this modified logarithmic barrier function is strongly convex in $\mathbf{x}$. Determine the strong convexity factor m.

(8 Marks)

Solution: The constraint $||x||^2 \le \rho^2$ adds the term

$-\mu \log(\rho^2 - \|x\|^2)$ to the logarithmic barrier

$$\Rightarrow \nabla^2(\bar{B}(x,\mu)) = \nabla^2(B(x,\mu)) + \frac{2\mu}{\rho^2 - \|x\|^2} I + \frac{4\mu}{(\rho^2 - \|x\|^2)^2} \|x\|^2$$

$$\succeq \nabla^2(B(x,\mu)) + \frac{2\mu}{\rho^2} I$$

$$\succeq \frac{2\mu}{\rho^2} I \quad (\because B(x,\mu) \text{ is convex})$$

$\therefore$ strong convexity factor $\underbrace{m = \frac{2\mu}{\rho^2}}_{\text{one possible value}}$

Frictional loss thru a horizontal pipeline

For

$L = 3000$ ft $\simeq 1$ km

$$H = \frac{4FLV^2}{2gD}$$

velocity needs to be estimated in ft/min

Diameter

Delivery head

9.8 m/s²

$$HP\ \text{reqd} = \frac{W\rho H}{75}$$

Flow rate of water m³/s . . . in terms of buckets per hr.

Density of water 1000 kg/m³

in m³/s

4000 litres per hr

= 1.2 litres/sec

$\simeq 1/2$ HP ← 37.5 kg m/s