The University of Texas at Austin
Department of Electrical and Computer Engineering

EE381V-11: Large Scale Optimization — Fall 2012

Problem Set One

Caramanis/Sanghavi — Due: Thursday, September 13, 2012.

This problem set reviews the theory we covered in the last few lectures, and also gets us thinking about descent algorithms. As in the previous problem set, problems marked with a "(?)" need not be turned in.

Reading Assignments

1. (?) Reading: Boyd & Vandenberghe: Chapters 9.1 & 9.2.

Matlab and Computational Assignments. Please provide a printout of the Matlab code you wrote to generate the solutions to the problems below.

1. This problem illustrates how the gradient descent algorithm behaves in different levels of strong convexity. To begin with, download the file: `http://users.ece.utexas.edu/~cmcaram/EE381V_2012F/ps1_matlab.zip`, which contains a matlab file that will generate the data for the problem.

   We have a simple unconstrained optimization problem:

   $$\min_{\beta \in \mathbb{R}^n} \quad f(\beta) \triangleq \frac{1}{2}\beta^T X \beta$$

   where $X \in \mathbb{R}^{n \times n}$ is a symmetric and positive definite matrix. In the matlab file, you can find three matrices for the problem, in which (a) all eigenvalues are one, (b) a half of the eigenvalues are one and the other half of them are very small, (c) all other than a few very large eigenvalues are one.

   We want to run the gradient descent algorithm which iteratively computes

   $$\beta^{(n+1)} = \beta^{(n)} - \gamma \nabla f(\beta^{(n)})$$

   where γ is a constant step size. The initial $\beta^{(0)}$ is the all-ones vector.

   For each matrix, find the range of γ that the solution converges to zero and the range of γ that the algorithm diverges, and explain why. Take example values of γ to illustrate the two behaviors, convergence to zero and divergence. Plot $f(\beta^{(n)})$ over n for the two of your values.

2. Take $\gamma = 1$, and plot $f(\beta^{(n)})$ over n for the second matrix (b) of the above three. Explain the convergence behavior of the solution based on the plot.

Written Problems

1. Various properties of orthogonal subspaces: Let V be a finite dimensional vector space with an inner produce, and let $U \subseteq V$ be a subspace. Recall that the space $U^{\perp}$ is defined as:

   $$U^{\perp} = \{v \in V \, : \, \langle v \, , \, u \rangle = 0, \ \forall u \in U\}.$$

(a) Show that if U is a subspace, then so is $U^\perp$.

(b) Show that $(U^\perp)^\perp = U$.

(c) Show that if $U, W \subseteq V$ are subspaces of V, then

$$U \subseteq W \;\Leftrightarrow U^\perp \supseteq W^\perp.$$

(d) Suppose now that $X \subseteq V$ is just a subset, i.e., not necessarily a subspace of V. Show that the definition $X^\perp$ still makes sense, and that $X^\perp$ is a subspace. Next show that $(X^\perp)^\perp \supseteq X$, and it is defined as the smallest subspace that contains the set X.

(e) Show that when U is a subspace of V, then V is the direct product of U and $U^\perp$ (denoted $V = U \oplus U^\perp$). That is, show that any $v \in V$ can be written *uniquely* as

$$v = u + u^\perp,$$

where $u \in U$, and $u^\perp \in U^\perp$.

2. (Boyd and Vandenberghe, Ex. 2.10) Consider the set

$$C = \{x \in \mathbb{R}^n \,:\, x^\top A x + b^\top x + c \leq 0\},$$

where $A \in \mathbb{S}^n$, $b \in \mathbb{R}^n$ and $c \in \mathbb{R}$.

(a) Show that if $A \in \mathbb{S}^n_+$ (i.e., A is positive semidefinite) then the set C is convex.

(b) Consider the set obtained by intersecting C with a hyperplane:

$$C_1 = C \cap \{x \,:\, g^\top x + h = 0\}.$$

Show that C_1 is convex if there exists $\lambda \in \mathbb{R}$ such that $(A + \lambda g g^\top) \in \mathbb{S}^n_+$.

3. (Boyd and Vandenberghe, Ex. 2.21) For $C, D \subseteq \mathbb{R}^n$ disjoint convex sets, let

$$\mathcal{S} = \{(a, b) \,:\, a^\top x \leq b \;\forall x \in C,\; a^\top x \geq b \;\forall x \in D\}$$

be the set of separating hyperplanes. Show that $\mathcal{S}$ is convex.

4. (?) In class we claimed that there are several natural operations on sets, that preserve convexity. Convince yourselves that the following all preserve convexity.

(a) Cartesian product: If $C_1, \ldots, C_m \subseteq \mathbb{R}^d$ are convex sets, then the set

$$C = C_1 \times \cdots \times C_m = \{(x_1, \ldots, x_m),\; x_i \in C_i\}$$

is convex.

(b) Affine and inverse maps: For $C \subseteq \mathbb{R}^n$ convex, and $A : \mathbb{R}^n \to \mathbb{R}^m$ a linear operator (i.e., an $m \times n$ matrix) then **show that** the following two sets are convex:

$$\begin{aligned} D_1 &= \{Ax \,:\, x \in C\} \\ D_2 &= \{x \,:\, Ax \in C\}. \end{aligned}$$

(c) Minkowski sum: If $C_1, C_2 \subseteq \mathbb{R}^n$ are convex, **show that**

$$C = C_1 + C_2 = \{x = x_1 + x_2 \,:\, x_1 \in C_1, x_2 \in C_2\}$$

is convex.

5. (?) Boyd and Vandenberghe, Ex. 2.26.
6. (?) Boyd and Vandenberghe, Ex. 2.35.
7. Consider two points, $v_1, v_2 \in \mathbb{R}^n$. Show that there exist $c \in \mathbb{R}^n$ and $d \in \mathbb{R}$ (and find them!) such that

$$\{x \, : \, ||x - v_1|| \leq ||x - v_2||\} = \{x \, : \, c^\top x \leq d\}.$$

   Thus, you are showing that the set of points in $\mathbb{R}^n$ that are closer to point v_1 than to point v_2, form a half-space.
8. Let A be an $n \times m$ real matrix, and B a $k \times m$ real matrix. Suppose that for every $x \in \mathbb{R}^m$, $Ax = 0$ only if $Bx = 0$, that is,

$$Ax = 0 \Rightarrow Bx = 0.$$

   Show that there exists a $k \times n$ real matrix C such that $CA = B$.