Homework 1

Convex Optimization 10-725/36-725

Due Tuesday September 13 at 5:30pm
submitted to Christoph Dann in Gates 8013
(Remember to a submit separate writeup for each problem, with your name at the top)

Total: 75 points

1 Convex sets (20 points)

(a, 12 pts) Closed sets and convex sets.

i. Show that a polyhedron $\{x \in \mathbb{R}^n : Ax \le b\}$, for some $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, is both convex and closed.

ii. Show that if $S_i \subseteq \mathbb{R}^n, i \in I$ is a collection of convex sets, then their intersection $\cap_{i\in I} S_i$ is also convex. Show that the same statement holds if we replace "convex" with "closed".

iii. Given an example of a closed set in $\mathbb{R}^2$ whose convex hull is not closed.

iv. Let $A \in \mathbb{R}^{m\times n}$. Show that if $S \subseteq \mathbb{R}^m$ is convex then so is $A^{-1}(S) = \{x \in \mathbb{R}^n : Ax \in S\}$, which is called the preimage of S under the map $A : \mathbb{R}^n \to \mathbb{R}^m$. Show that the same statement holds if we replace "convex" with "closed".

v. Let $A \in \mathbb{R}^{m\times n}$. Show that if $S \subseteq \mathbb{R}^n$ is convex then so is $A(S) = \{Ax : x \in S\}$, called the image of S under A.

vi. Give an example of a matrix $A \in \mathbb{R}^{m\times n}$ and a set $S \subseteq \mathbb{R}^n$ that is closed and convex but such that $A(S)$ is not closed.

(b, 4 pts) The following is an important property of polyhedra:

$$P \subseteq \mathbb{R}^{m+n} \text{ is a polyhedron } \Rightarrow \{x \in \mathbb{R}^n : (x, y) \in P \text{ for some } y \in \mathbb{R}^m\} \text{ is a polyhedron.} \quad (1)$$

(Bonus: prove this property.)

i. Use the above property (1) about polyhedra to conclude that if $A \in \mathbb{R}^{m\times n}$ and $P \subseteq \mathbb{R}^n$ is a polyhedron then $A(P)$ is a polyhedron.

ii. Give an example to show that (1) is no longer true if we replace "polyhedron" with "closed and convex set".

(c, 4 pts) The following is a "strict" variant of the Separating Hyperplane Theorem: if $C, D \subseteq \mathbb{R}^n$ are disjoint, closed and convex, and (say) D is bounded, then there exists $a \in \mathbb{R}^n, b \in \mathbb{R}$ with $a \ne 0$ such that $a^T x > b$ for all $x \in C$ and $a^T x < b$ for all $x \in D$ (i.e., the hyperplane $\{x \in \mathbb{R}^n : a^T x = b\}$ strictly separates C, D). Use this to prove *Farkas' Lemma*: given $A \in \mathbb{R}^{m\times n}, b \in \mathbb{R}^m$, exactly one of the following is true:

- $\exists\, x \in \mathbb{R}^n$ such that $Ax = b, x \geq 0$;
- $\exists\, y \in \mathbb{R}^m$ such that $A^T y \geq 0, y^T b < 0$.

(Hint: it will help you to use part (b.i), to deduce that the set $\{Ax : x \geq 0\}$ is a polyhedron, and hence closed and convex by part (a.i).)

2 Convex functions (16 points)

(a, 6 pts) Prove that $f(x, y) = |xy| + a(x^2 + y^2)$ is convex if and only if $a \geq 1/2$. Also prove that it is strongly convex if $a > 1/2$. (Bonus: produce 3d plots of f, one for each $a \in \{0, 1/4, 1/2, 3/4\}$.)

(b, 6 pts) In each case below specify whether the function is strongly convex, strictly convex, convex, or nonconvex, and give a brief justification.

i. The *logarithmic barrier function*, $f : \mathbb{R}^n_{++} \to \mathbb{R}$ defined as

$$f(x) = -\sum_{i=1}^{n} \log(x_i).$$

ii. The *entropy function*, $f : \{x \in \mathbb{R}^n_+ : \sum_{i=1}^n x_i = 1\} \to \mathbb{R}$ defined as

$$f(x) = \begin{cases} -\sum_{i=1}^n x_i \log(x_i) & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

(c, 4 points) Let f be twice differentiable, with $\mathrm{dom}(f)$ convex. Prove that f is convex if and only if

$$(\nabla f(x) - \nabla f(y))^T (x - y) \geq 0,$$

for all x, y. This property is called *monotonicity* of the gradient ∇f.

3 Lipschitz gradients and strong convexity (16 points)

Let f be convex and twice differentiable.

(a, 8 pts) Show that the following statements are equivalent.

i. ∇f is Lipschitz with constant L;

ii. $(\nabla f(x) - \nabla f(y))^T (x - y) \leq L\|x - y\|_2^2$ for all x, y;

iii. $\nabla^2 f(x) \preceq LI$ for all x;

iv. $f(y) \leq f(x) + \nabla f(x)^T (y - x) + \frac{L}{2}\|y - x\|_2^2$ for all x, y.

(b, 8 pts) Show that the following statements are equivalent.

i. f is strongly convex with constant m;

ii. $(\nabla f(x) - \nabla f(y))^T (x - y) \geq m\|x - y\|_2^2$ for all x, y;

iii. $\nabla^2 f(x) \succeq mI$ for all x;

iv. $f(y) \geq f(x) + \nabla f(x)^T (y - x) + \frac{m}{2}\|y - x\|_2^2$ for all x, y.

4 Solving optimization problems with CVX (23 points)

CVX is a fantastic framework for disciplined convex programming—it's rarely the fastest tool for the job, but it's widely applicable, and so it's a great tool to be comfortable with. In this exercise we will set up the CVX environment and solve a convex optimization problem.

In this class, your solution to coding problems should include plots and whatever explanation necessary to answer the questions asked. In addition, full code should be submitted as an appendix to the homework document.

CVX variants are available for each of the major numerical programming languages. There are some minor syntactic and functional differences between the variants but all provide essentially the same functionality. The Matlab version (and by extension, the R version which calls Matlab under the covers) is the most mature but all should be sufficient for the purposes of this class.

Download the CVX variant of your choosing:

- Matlab - `http://cvxr.com/cvx/`
- Python - `http://www.cvxpy.org/en/latest/`
- Julia - `https://github.com/JuliaOpt/Convex.jl`
- R - `http://faculty.bscb.cornell.edu/~bien/cvxfromr.html`

and consult the documentation to understand the basic functionality. Make sure that you can solve the least squares problem $\min_\theta \ \|y - X\theta\|_2^2$ for a vector y and matrix X. Check your answer by comparing with the analytic least squares solution.

(a) Using CVX, we will solve the 2d fused lasso problem discussed in Lecture 1:

$$\min_{\theta} \ \frac{1}{2}\sum_{i=1}^{n}(y_i - \theta_i)^2 + \lambda \sum_{\{i,j\}\in E} |\theta_i - \theta_j|.$$

The set E is the set of all undirected edges connecting horizontally or vertically neighboring pxiels in the image. More specifically, $\{i,j\} \in E$ if and only if pixel i is the immediate neighbor of pixel j on the left, right, above or below.

1. (9 pts) Load the basic test data from `circle.csv` and solve the 2d fused lasso problem with $\lambda = 1$. Report the objective value obtained at the solution and plot the solution and original data as images. Why does the shape change its form?
2. (6 pts) Next, we consider how the solution changes as we vary λ. Load a grayscale 64×64 pixel version of the standard Lenna test image from `lenna_64.csv` and solve the 2d fused lasso problem for this image for $\lambda \in \{10^{-k/4} \ : \ k = 0, 1, \ldots, 8\}$. For each λ, report the value of the optimal objective value, plot the optimal image and show a histogram of the pixel values (100 bins between values 0 and 1). What change in the histograms can you observe with varying λ?

(b, 8 pts) Disciplined convex programming or DCP is a system for composing functions while ensuring their convexity. It is the language that underlies CVX. Essentially, each node in the parse tree for a convex expression is tagged with attributes for curvature (convex, concave, affine, constant) and sign (positive, negative) allowing for reasoning about the convexity of entire expressions. The website `http://dcp.stanford.edu/` provides visualization and analysis of simple expressions.

Typically, writing problems in the DCP form is natural, but in some cases manipulation is required to construct expressions that satisfy the rules. For each set of mathematical expressions below (all define a convex set), give equivalent DCP expressions along with a brief explanation of why the DCP expressions are equivalent to the original. DCP expressions should be given in a form that passes analysis at `http://dcp.stanford.edu/analyzer`.

Note: this question is really about developing a better understanding of the various composition rules for convex functions.

1. $\|(x, y)\|_1^3 \leq 5x + 7$
2. $\frac{2}{x} + \frac{9}{z-y} \leq 3, x > 0, y < z$
3. $\sqrt{x^2 + 4} + 2y \leq -5x$
4. $(x + 3)z(y - 5) \geq 8, x \geq -3, z \geq 0, y \geq 5$
5. $\frac{(x+3z)^2}{\log y} + 2y^2 \leq 10, y > 1$
6. $\log\left(e^{-\sqrt{x}} + e^{2z}\right) \leq -e^{5y}, x \geq 0$
7. $\sqrt{\|(2x - 3y, y + x)\|_1} = 0$
8. $y \log\left(\frac{y}{2x}\right) \leq y + x - 30, x > 0, y > 0$