Homework 6

Jordan Saethre

November 5, 2020

1. Beck Exercise 7.5. Let $f\colon \mathbb{R}^n \to \mathbb{R}$. Prove that f is convex if and only if for any $x \in \mathbb{R}^n$ and $d \neq 0$, the one-dimensional function $g_{x,d}(t) = f(x + td)$ is convex.

   Proof. ($\implies$)

$$\begin{aligned}
g(\lambda t + (1-\lambda)s) &= f(x + (\lambda t + (1-\lambda)s)d) \\
&= f(x + \lambda td + (1-\lambda)sd) \\
&= f(\lambda x + (1-\lambda)x + \lambda td + (1-\lambda)sd) \\
&= f(\lambda(x + td) + (1-\lambda)(x + sd)) \\
&\leq \lambda f(x + td) + (1-\lambda)f(x + sd) \\
&= \lambda g(t)_((1-\lambda)g(s)
\end{aligned}$$

   ($\impliedby$) Reverse the algebra steps used to show the forward implication:

$$\begin{aligned}
f(\lambda(x + td) + (1-\lambda)(x + sd)) &= f(\lambda x + (1-\lambda)x + \lambda td + (1-\lambda)sd) \\
&= f(x + (\lambda t + (1-\lambda)s)d) \\
&= g(\lambda t + (1-\lambda)s) \\
&\leq \lambda g(t)_((1-\lambda)g(s) \\
&= \lambda f(x + td) + (1-\lambda)f(x + sd)
\end{aligned}$$

   □

2. Beck Exercise 7.26. Let C be a convex subset of $\mathbb{R}^n$. A function f is called *strongly convex* over C if there exists $\sigma > 0$ such that the function $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex over C. The parameter σ is called the *strong convexity parameter*. In the following questions C is a given convex subset of $\mathbb{R}^n$.

   (i) Prove that f is strongly convex over C with parameter σ if and only if

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\sigma}{2}\lambda(1-\lambda)\|x - y\|^2$$

   for any $x, y \in C$ and $\lambda \in [0, 1]$.

Proof. ($\Longrightarrow$) Assume f is strongly convex, that is

$$f(\lambda x+(1-\lambda)y)-\frac{\sigma}{2}(\lambda x+(1-\lambda)y)^t(\lambda x+(1-\lambda)y) \leq \lambda(f(x)-\frac{\sigma}{2}x^tx)+(1-\lambda)(f(y)-\frac{\sigma}{2}y^ty)$$

Reorganizing terms we get

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\sigma}{2}\lambda(1-\lambda)\left[x^tx + 2y^tx + y^ty\right]$$

The last term in the square brackets is just $\|x-y\|^2$. Therefore

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\sigma}{2}\lambda(1-\lambda)\|x-y\|^2$$

($\Longleftarrow$) Assume the following inequality holds

$$f(\lambda x + (1-\lambda)y) \leq \lambda f(x) + (1-\lambda)f(y) - \frac{\sigma}{2}\lambda(1-\lambda)\|x-y\|^2$$

Now reverse the algebra steps that were done to show the forward implication.

□

(ii) Prove that a strongly convex function over C is also strictly convex over C.

Proof. Assume $f(x) - \frac{\sigma}{2}\|x\|^2$ is convex.

$$f(\lambda x+(1-\lambda)y)-\frac{\sigma}{2}\|\lambda x+(1-\lambda)y\|^2 \leq \lambda\left(f(x)-\frac{\sigma}{2}\|x\|^2\right)+(1-\lambda)\left(f(x)-\frac{\sigma}{2}\|x\|^2\right)$$

After reorganizing terms we have

$$f(\lambda x+(1-\lambda)y) \leq \lambda f(x)+(1-\lambda)f(y)-\frac{\sigma}{2}\left[\lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \|\lambda x + (1-\lambda)y\|^2\right]$$

Notice that since $\|\cdot\|^2$ is convex we have

$$0 \leq \lambda\|x\|^2 + (1-\lambda)\|y\|^2 - \|\lambda x + (1-\lambda)y\|^2$$

Therefore

$$f(\lambda x + (1-\lambda)y) < \lambda f(x) + (1-\lambda)f(y)$$

□

(iii) Suppose that f is continuously differentiable over C. Prove that f is strongly convex over C with parameter σ if and only if

$$f(y) \geq f(x) + \nabla f(x)^t(y-x) + \frac{\sigma}{2}\|x-y\|^2$$

for any $x, y \in C$.

Proof. ($\Longrightarrow$) Assume that f is strongly convex. Then by 2(i) from above we can write

$$f(\lambda y + (1-\lambda)x) \leq \lambda f(y) + (1-\lambda)f(x) - \frac{\sigma}{2}\lambda(1-\lambda)\|y-x\|^2$$

Reorganize terms to obtain

$$\frac{f(\lambda y + (1-\lambda)x) - f(x)}{\lambda} \leq f(y) - f(x) - \frac{\sigma}{2}(1-\lambda)\|y-x\|^2$$

Now let $\lambda \to 0^+$ and the left hand side becomes $\nabla f(x)^t(y-x)$ and the last term on the right hand side becomes just $\frac{\sigma}{2}\|y-x\|^2$. Therefore

$$f(y) \geq f(x) + \nabla f(x)^t(y-x) + \frac{\sigma}{2}\|x-y\|^2$$

($\Longleftarrow$) Assume the inequality above holds and we will show that f is strongly convex. Let $z, w \in C$ and $\lambda \in (0,1)$. Let $u = \lambda z + (1-\lambda)w$, then

$$z - u = \frac{u-(1-\lambda)w}{\lambda} - u = \frac{-1-\lambda}{\lambda}(w-u)$$

Now using the inequality above we can write

$$f(z) \geq f(u) + \nabla f(u)^t(z-u) + \frac{\sigma}{2}\|u-z\|^2 \tag{1}$$

$$f(w) \geq f(u) - \frac{\lambda}{1-\lambda}\nabla f(u)^t(z-u) + \frac{\lambda^2}{(1-\lambda)^2}\frac{\sigma}{2}\|u-z\|^2 \tag{2}$$

Now multiply inequality (1) by $\frac{\lambda}{1-\lambda}$ and add to (2) to get

$$\frac{\lambda}{1-\lambda}f(z) + f(w) \geq \frac{\lambda}{1-\lambda}f(u) + f(u) + \frac{\lambda}{(1-\lambda)^2}\frac{\sigma}{2}\|u-z\|^2$$

Multiply through by $1-\lambda$

$$f(u) \leq \lambda f(z) + (1-\lambda)f(w) - \frac{\lambda}{1-\lambda}\frac{\sigma}{2}\|u-z\|^2$$

Notice that $z - u = (1-\lambda)(z-w)$

$$f(\lambda z + (1-\lambda w)) \leq \lambda f(z) + (1-\lambda)f(w) - \lambda(1-\lambda)\frac{\sigma}{2}\|z-w\|^2$$

Which by 2 (i) above implies that f is strongly convex. $\square$

(iv) Suppose that f is continuously differentiable over C. Prove that f is strongly convex over C with parameter σ if and only if

$$(\nabla f(x) - \nabla f(y))^t (x - y) \geq \sigma \|x - y\|^2$$

for any $x, y \in C$.

Proof. ($\Longrightarrow$) By 2 (iii) above we can write

$$f(y) \geq f(x) + \nabla f(x)^t (y - x) + \frac{\sigma}{2}\|x - y\|^2$$
$$f(x) \geq f(y) + \nabla f(y)^t (x - y) + \frac{\sigma}{2}\|y - x\|^2$$

Simply add these two inequalities together to obtain

$$(\nabla f(x) - \nabla f(y))^t (x - y) \geq \sigma \|x - y\|^2$$

($\Longleftarrow$) Assume the last inequality above and consider $g(t) = f(x + t(y - x))$ where $t \in [0, 1]$. By the Fundamental Theorem of Calculus we have

$$\begin{aligned} f(y) = g(1) = g(0) &+ \int_0^1 g'(t)dt \\ &= f(x) + \int_0^1 (y - x)^t \nabla f(x + t(y - x))dt \\ &= f(x) + \nabla f(x)^t (y - x) + \int_0^1 (y - x)^t (\nabla f(x + t(y - x)) - \nabla f(x))dt \end{aligned}$$

The term inside the integral is greater than or equal to $\sigma\|x - y\|^2$ by our original assumption. Therefore

$$f(y) \geq f(x) + \nabla f(x)^t (y - x) + \frac{\sigma}{2}\|x - y\|^2$$

Which by 2 (iii) implies that f is strongly convex. $\square$

(v) Suppose that f is twice continuously differentiable over C. Prove that f is strongly convex over C with parameter σ if and only if $\nabla^2 f(x) \succeq \sigma I$ for any $x \in C$.

Proof. ($\Longrightarrow$) Assume f is strongly convex and let $x \in C$ and $y \in \mathbb{R}$ then $x + \lambda y \in C$ when $\lambda \in (0, \epsilon)$ where $\epsilon > 0$ is sufficiently small. By 2 (iii) we can write

$$f(x + \lambda y) \geq f(x) + \lambda \nabla f(x)^t y + \frac{\sigma}{2}\lambda^2 \|y\|^2$$

The left hand side of this inequality by the Quadratic approximation theorem is

$$f(x + \lambda y) = f(x)\lambda \nabla f(x)^t y + \frac{\lambda^2}{2} y^t \nabla^2 f(x) y + o(\lambda^2 \|y\|^2)$$

Combining these two facts and canceling terms we find that

$$\frac{\lambda^2}{2} y^t \nabla^2 f(x) y + o(\lambda^2 \|y\|^2) - \frac{\sigma}{2} \lambda^2 \|y\|^2 \geq 0$$

If we divide by λ^2 and let $\lambda \to 0^+$ the little o term will go to zero and we have

$$\frac{1}{2} y^t \nabla^2 f(x) y - \frac{\sigma}{2} \|y\|^2 \geq 0$$

This can further be rewritten as

$$y^t \left(\nabla^2 f(x) - \sigma I\right) y \geq 0$$

Which shows that $\nabla^2 f(x) \succeq \sigma I$.
($\Longleftarrow$) Let $x, y \in C$ then by the Linear Approximation Theorem there exists $z \in [x, y]$ (which is in C) where

$$f(y) = f(x) + \nabla f(x)^t (y - x) \frac{1}{2} (y - x)^t \nabla^2 f(z) (y - x)$$

Since $\nabla^2 f(x) \succeq \sigma I$ we can write

$$\frac{1}{2} (y - x)^t (\nabla^2 f(z) - \sigma I)(y - x) \geq 0$$

Which implies

$$\frac{1}{2} (y - x)^t \nabla^2 f(z) (y - x) \geq (y - x)^t \frac{\sigma}{2} I (y - x) = \frac{\sigma}{2} \|x - y\|^2$$

By 2 (iii) this implies f is strongly convex.

$\square$

3. Beck Exercise 8.1. Consider the problem

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & g(x) \leq 0 \\ & x \in X \end{aligned}$$

where f and g are convex functions over $\mathbb{R}^n$ and $X \subset \mathbb{R}^n$ is a convex set. Suppose that $x^\star$ is an optimal solution to this problem that satisfies $g(x^\star) < 0$. Show that $x^\star$ is also an optimal solution of the problem

$$\begin{aligned} \min\ & f(x) \\ \text{s.t. } & x \in X. \end{aligned}$$

Proof. (Proof by Contradiction) Assume that x^* is optimal for the first problem, but assume there is a more optimal point for the second problem, say $\tilde{x}$. This would imply that $f(\tilde{x}) < f(x^*)$. Let $\lambda \in (0, 1]$, there exists a point $z = \lambda\tilde{x} + (1-\lambda)x^*$ such that $g(z) < 0$. Since g is convex we can write

$$g(z) \leq \lambda g(\tilde{x}) + (1-\lambda)g(x^*)$$

But since $g(x^*) < 0$ and $g(z) < 0$ this implies that $g(\tilde{x}) < 0$ which contradicts the optimality of x^* for the first problem. Therefore if x^* is optimal for the first problem then it is also optimal for the second problem. $\square$

4. Beck Exercise 8.2. Let $C = B[x_0, r]$, where $x_0 \in \mathbb{R}^n$ and $r > 0$ are given. Find a formula for the orthogonal projection operator P_C.

   The formula for the orthogonal projection operator P_C is the solution to the following optimization problem:
   $$\min_y\{\|y - x\|^2 : \|y - x_0\|^2 \leq r^2\}$$
   If $\|x - x_0\|^2 \leq r$ then x is in C and the optimal solution is $y = x$. If $\|x - x_0\|^2 > r$ then x is not in C and the closest point $y \in C$ will exist on the boundary of C. Therefore this problem is equivalent to

   $$\min_y\{\|y - x\|^2 : \|y - x_0\|^2 = r^2\}$$

   Now consider the line passing through x and x_0 given by

   $$y = t(x - x_0) + x_0 \quad \forall t \in \mathbb{R}$$

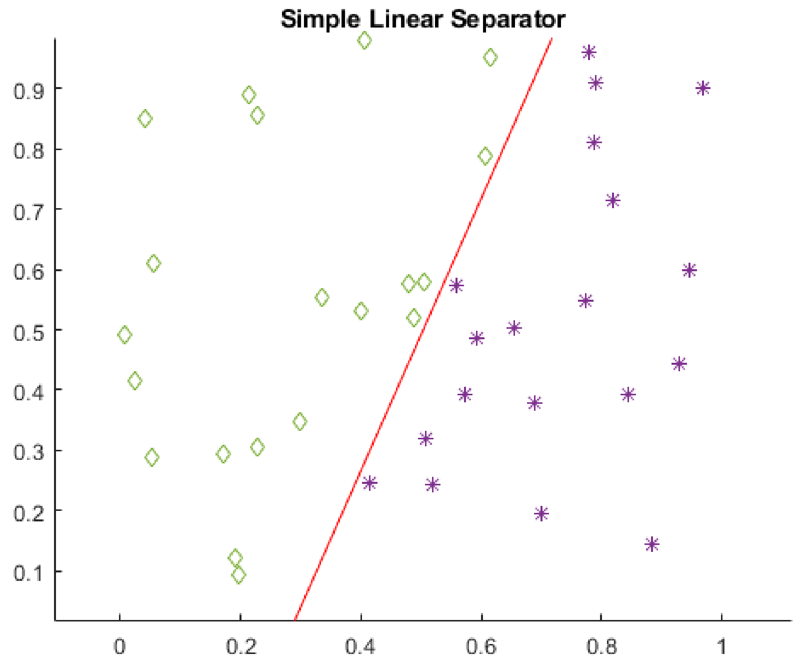
   The point where this line intersects the boundary of C is the optimal solution to the problem above.

   $$\begin{gathered} y = t(x - x_0) + x_0 \\ \implies t = \frac{\|y - x_0\|}{\|x - x_0\|} = \frac{r}{\|x - x_0\|} \end{gathered}$$

   Therefore $y = \frac{r(x-x_0)}{\|x-x_0\|} + x_0$ and the formula for P_C is

   $$P_C(x) = \begin{cases} x & if \ \ \|x - x_0\| \leq r \\ \frac{r(x-x_0)}{\|x-x_0\|} + x_0 & if \ \ \|x - x_0\| > r \end{cases}$$

5. Beck Exercise 8.5. Suppose that we are given 40 points in the plane. Each of these points belongs to one of two classes. Specifically, there are 19 points of class 1 and 21 points of class 2. The plot of the points is given in Figure 8.8. Note that the rows of $A_1 \in \mathbb{R}^{19\times 2}$ are the 19 points of class 1 and the rows of $A_2 \in \mathbb{R}^{21\times 2}$ are the 21 points of class 2. Write a CVX-based code for finding the maximum-margin line separating the two classes of points.




```
% (Code) simple linear separator

% data generation
rand('seed',314)
x = rand(40,1);   y = rand(40,1);
class = [2*x < y + 0.5] +1
A1 = [x(find(class == 1)), y(find(class==1))];
A2 = [x(find(class == 2)), y(find(class==2))];

% plot classes 1 and 2
figure(1); hold on;
plot(A1(:,1), A1(:,2), '*', 'MarkerSize', 6);
plot(A2(:,1), A2(:,2), 'd', 'MarkerSize', 6);

% solve the qp
cvx_begin
variables w(2) b(1)
A1*w - b >= 1;
A2*w- b <= -1;
```

```
minimize(norm(w));
cvx_end

% plot the solution
t_min = min([A1(1,:),A2(1,:)]);
t_max = max([A1(1,:),A2(1,:)]);
t = linspace(t_min-1,t_max+1,100);
p = -w(1)*t/w(2) + b/w(2);
plot(t,p, '-r');
axis equal
title('Simple Linear Separator');
```