Homework 7

Wednesday, November 18, 2020 07:16

Jordan Saethre

1) Consider the problem

$$\min\{f(x) : x \geq 0 , e^t x = 1\}$$

Where f is continuously differentiable over the $\{x \mid x \geq 0, e^t x = 1\}$. Prove that x^* is a stationary point if and only if

$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} = \mu & x_i^* > 0 \\ \geq \mu & x_i^* = 0 \end{cases}$$

Proof:

$(\Rightarrow)$ $\nabla f(x^*)^t (x - x^*) \geq 0 \quad \forall \; x$ s.t. $e^t x = 1$ $x \geq 0$

$$\Rightarrow \nabla f(x^*)^t x - \nabla f(x^*)^t x^* \geq 0$$

$$\Rightarrow \sum_{i=1}^{n} \frac{\partial f}{\partial x_i^*} x_i - \sum_{i=1}^{n} \frac{\partial f}{\partial x_i^*} x_i^* \geq 0$$

$$\Rightarrow \sum_{i=1}^{n} \frac{\partial f}{\partial x_i^*} x_i \geq \sum_{i=1} \frac{\partial f}{\partial x_i^*} x_i^*$$

Since $\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i^* = 1$ this implies

the last inequality can only hold if

the value of each component of $\nabla f(x^*)$

is either $\geq \mu \in \mathbb{R}$ for $x_i^* = 0$ or

$= \mu$ if $x_i^* > 0$.

($\Leftarrow$) For the other direction assume

$$\frac{\partial f}{\partial x_i}(x^*) \begin{cases} = \mu & x_i^* > 0 \\ \geq \mu & x_i^* = 0 \end{cases}$$

If $x_i^* > 0$ $\forall i$ and $e^t x^* = e^t x = 1$ and
Let $\frac{\partial f}{\partial x_i^*} = \mu$ $\forall i$, then

$$\mu \sum_{i=1}^{n} x_i - \mu \sum_{i=1}^{n} x_i^* = \mu - \mu = 0$$

If $x_j^* = 1$ and $x_i^* = 0$ $\forall i \neq j$ and
let $\frac{\partial f}{\partial x_i^*} = \lambda \geq \mu$ $\forall i \neq j$ then

Without loss of generality assume $j=1$

$$\underbrace{\lambda x_1}_{\leq \lambda} + \underbrace{\sum_{i=2}^{n} \mu x_i}_{\leq \mu} + \underbrace{\lambda x_1^*}_{= \lambda} + \underbrace{\sum_{i=2}^{n} \mu x_i^*}_{= 0}$$

$\Rightarrow \quad \geq 0$

$\square$

2) Let f be a strongly convex function over a closed convex set $S \subseteq \mathbb{R}^n$ w/ strong covexity parameter σ, i.e.

$$f(y) \geq f(x) + \nabla f(x)^t (y-x) + \frac{\sigma}{2} \|x-y\|^2 \quad \forall x,y \in S$$

Assume $f \in C_L^{1,1}(S)$.

Assume $f \in C_L^{1,1}(S)$.

Consider $\min\{f(x) : x \in S\}$ where S is a closed convex set. Let $\{x_k\}_{k\geq 0}$ be the sequence generated by the gradient projection method for solving the problem w/ constant step size $t_k = \frac{1}{L}$.

Let x^* be the optimal solution and f^* the optimal value.

Prove $\exists\ c \in (0,1)$ s.t. for $k \geq 0$

$$\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|$$

Find an explicit expression for c.

Proof:

Since $f \in C_L^{1,1}(S)$ we can use the descent lemma:

$$f(x_{k+1}) \leq f(x_k) + \langle \nabla f(x_k), x_{k+1} - x_k \rangle + \frac{L}{2}\|x_k - x_{k+1}\|^2$$

Also, since f is strongly convex:

$$\frac{\sigma}{2}\|x_k - x^*\| + f(x_k) + \nabla f(x_k)^t(x^* - x_k) \leq f(x^*)$$

$$\Rightarrow \quad f(x_k) \leq f(x^*) - \nabla f(x_k)^t(x^* - x_k) - \frac{\sigma}{2}\|x_k - x^*\|^2$$

$$\Rightarrow \quad f(x_k) \leq f(x^*) + \nabla f(x_k)^t(x_k - x^*) - \frac{\sigma}{2}\|x_k - x^*\|^2$$

Now plugging these into Thm 9.16

Now plugging these into Thm 9.16

$$\frac{2}{L}\left(f(x_{k+1}) - f(x^*)\right) \leq \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2$$

$$\frac{2}{L}\left[\cancel{f(x^*)} + \nabla f(x_k)^t(x_k - x^*) - \frac{\sigma}{2}\|x_k - x^*\|^2 - \cancel{f(x^*)}\right] \leq \|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2$$

$$\frac{2}{L}\nabla f(x_k)^t(x_k - x^*) \leq \|x_k - x^*\|^2 + \frac{\sigma}{L}\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2$$

$$\frac{2}{L}\nabla f(x_k)^t(x_k - x^*) \leq \left(1 + \frac{\sigma}{L}\right)\|x_k - x^*\|^2 - \|x_{k+1} - x^*\|^2$$

Now since x^* is an optimal point this implies it is stationary and hence the RHS is ≥ 0 therefore:

$$\|x_{k+1} - x^*\| \leq \sqrt{1 + \frac{\sigma}{L}}\,\|x - x^*\|$$

$\square$

3) Consider the problem

$$\min\ x_1^2 + 2x_2^2 + x_1$$

$$\text{s.t. } x_1 + x_2 \leq a \qquad \text{for } a \in \mathbb{R}$$

(i) Prove $\forall a \in \mathbb{R}$ the problem has a unique optimal solution.

Proof: clearly as $\|x\| \to \infty$ the function increases without bound and is by definition coercive. Now by thm 2.32, attainment under coerciveness, we have a continuous coercive function defined on a attainment under coerciveness, we have a continuous coercive function defined on a nonempty closed space, in this case the half space $x_1 + x_2 \leq a$, which implies that f has a global minimum over this half space.

However, this does not imply that the global minimum is unique. For this we will recall that if f is strictly convex then there exists at most one optimal solution, (Thm 8.3). Therefore to prove the desired we need only show that f is strictly convex.

f can be written in quadratic form as

$$f(x_1, x_2) = [x_1 \ \ x_2]\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [1 \ \ 0]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

$$= [x_1 \ \ 2x_2]\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + x_1$$

$$= x_1^2 + 2x_2^2 + x_1$$

Thm 7.10 says that for a quadratic function

$$f(x) = x^t A x + 2b^t x + c$$

if $A \succ 0$ then f is strictly convex which since $A = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 2 \end{smallmatrix}\right]$ is diagonal with all strictly positive diagonal entries this implies $A \succ 0$ (lemma 2.19).

Therefore f is strictly convex and there exists at most one optimal solution.

Therefore f is strictly convex and there exists at most one optimal solution.

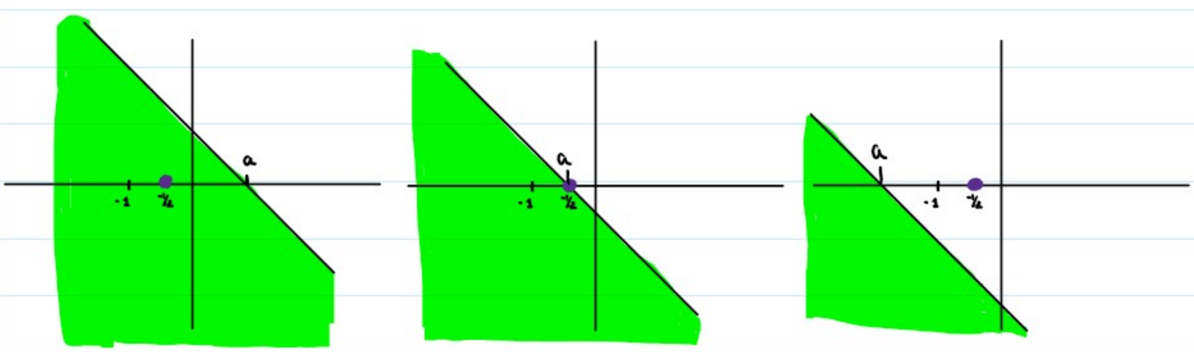
(ii) Solve the problem in terms of a.

Since we know that there exists a unique global min lets first consider what it would be for the associated unconstrained problem.

$$\nabla f(x) = \begin{bmatrix} 2x_1 + 1 \\ 4x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \Rightarrow \begin{matrix} x_1 = -\frac{1}{2} \\ x_2 = 0 \end{matrix}$$

$x^* = (-\frac{1}{2}, 0)$ is the only critical point which is what we expected and we know by the argument above in (i) that it is a global minimum.

Now consider what happens taking into account the constraint $x_1 + x_2 \leq a$:



$a > -\frac{1}{2}$ $\qquad\qquad$ $a = -\frac{1}{2}$ $\qquad\qquad$ $a < -\frac{1}{2}$

For $a \geq -\frac{1}{2}$ the optimal point is $(-\frac{1}{2}, 0)$, but when $a < -\frac{1}{2}$ the optimal point will be on when $a < -\frac{1}{2}$ the optimal point will be on the boundary $x_1 + x_2 = a$. We can solve for x_2 and plug into the original function. We now have a simple 1D minimization problem.

$$x_2 = a - x_1$$

$$f(x_1) = x_1^2 + 2(a - x_1)^2 + x_1$$

$$= 3x_1^2 + (1-4a)x_1 + 2a^2$$

$$f'(x_1) = 6x_1 + 1 - 4a = 0$$

$$\Rightarrow x_1 = \frac{4a-1}{6}$$

$$\Rightarrow x_2 = \frac{2a+1}{6}$$

$\therefore$ the optimal solution for $a < -\frac{1}{2}$

$$x^* = \left(\frac{4a-1}{6}, \frac{2a+1}{6}\right)$$

(iii) Let $f(a)$ be the optimal value of the problem with parameter a. Write an explicit expression for f and prove it is a convex function.

$$f(x_1, x_2) = x_1^2 + 2x_2^2 + x_1$$

$$f(-\tfrac{1}{2}, 0) = (-\tfrac{1}{2})^2 + 2(0) + (-\tfrac{1}{2}) = -\tfrac{1}{4}$$

$$f(a) = \begin{cases} \frac{-1}{4} & \text{for } a \geq -\frac{1}{2} \\ \left(\frac{4a-1}{6}\right)^2 + 2\left(\frac{2a+1}{6}\right)^2 + \left(\frac{4a-1}{6}\right) & \text{for } a < -\frac{1}{2} \end{cases}$$

$f(a)$ is convex if

$$f(\lambda a + (1-\lambda)b) \leq \lambda f(a) + (1-\lambda) f(b) \quad \forall\, a, b \in \mathbb{R}$$
$$\lambda \in [0,1]$$

First simplify $f(a)$

$$f(a) = \begin{cases} -1/4 & \text{for } a \geq -\frac{1}{2} \\ \frac{2}{3}a^2 + \frac{32}{9}a - \frac{33}{36} & \text{for } a < -\frac{1}{2} \end{cases}$$

$$f''(a) = \begin{cases} 0 & \text{for } a \geq -\frac{1}{2} \\ 4/3 & \text{for } a < -\frac{1}{2} \end{cases}$$

Both cases have a nonnegative second derivative therefore $f(a)$ is convex.