MATH60005/70005: Optimisation (Autumn 24-25)

Chapter 6: solutions

Dr Dante Kalise, Dr Estefanía Loayza-Romero and Sara Bicego
Department of Mathematics
Imperial College London, United Kingdom
{dkaliseb,kloayzar,s.bicego21}@imperial.ac.uk

1. Derive the orthogonal projection formula for a closed ball centered at $\mathbf{x}_0 \in \mathbb{R}^n$, $B[\mathbf{x}_0, r]$.
2. Show that the stationarity condition over the unit ball in $\mathbb{R}^n$, that is,

$$\min\{f(\mathbf{x}) : \|\mathbf{x}\| \leq 1\}$$

is given by $\nabla f(\mathbf{x}^*) = 0$, or $\|\mathbf{x}^*\| = 1$ and there exists $\lambda \leq 0$ such that $\nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$.

3. Consider the minimization problem

$$\begin{aligned} \min \quad & 2x_1^2 + 3x_2^2 + 4x_3^2 + 2x_1x_2 - 2x_1x_3 - 8x_1 - 4x_2 - 2x_3 \\ \text{subject to} \quad & x_1, x_2, x_3 \geq 0 \,. \end{aligned}$$

- Show that the vector $(\frac{17}{7}, 0, \frac{6}{7})^\top$ is an optimal solution.
- Implement a projected gradient method with constant stepsize $\frac{1}{L}$, where L is the Lipschitz constant of the gradient of the function.

Solutions

1. We need to solve the problem

$$\min_{\mathbf{y}} \|\mathbf{y} - \mathbf{x}\|^2 \qquad \text{subject to} \qquad \|\mathbf{y} - \mathbf{x}_0\| \leq r \,.$$

Using the change of variables $\mathbf{z} = \mathbf{x} - \mathbf{x}_0$, then, then problem can be rewritten as

$$\min_{\mathbf{z}} \|\mathbf{z} - (\mathbf{x} - \mathbf{x}_0)\|^2 \qquad \text{subject to} \qquad \|\mathbf{z}\| \leq r \,,$$

Notice that the optimal solution of the last problem is $\mathbb{P}_{B[\mathbf{0},r]}(\mathbf{x} - \mathbf{x}_0)$. Thus,

$$\mathbb{P}_{B[\mathbf{x}_0,r]} = \mathbf{x}_0 + \mathbb{P}_{B[\mathbf{0},r]}(\mathbf{x} - \mathbf{x}_0)$$

Therefore, we only need to derive an expression for $\mathbb{P}_{B[\mathbf{0},r]}(\mathbf{x})$. This problem can be written as

$$\min_{\mathbf{y}} \|\mathbf{y}-\mathbf{x}\|^2 \qquad \text{subject to} \qquad \|\mathbf{y}\|^2 \leq r^2\,,$$

which in turn – by expanding the square and $\|y\| \leq r \iff \|\mathbf{y}\|^2 \leq r^2$ – becomes

$$\min_{\mathbf{y}} \|\mathbf{y}\|^2 + \|\mathbf{x}\|^2 - 2\mathbf{x}^\top \mathbf{y} \qquad \text{subject to} \qquad \|\mathbf{y}\|^2 \leq r^2\,.$$

If $\|\mathbf{x}\| \leq r$, then $\mathbb{P}_{B[\mathbf{0},r]} = \mathbf{x}$. On the other hand, when $\|x\| > r$ we know that $\mathbf{y}$ lies in the boundary, hence $\|\mathbf{y}\|^2 = r^2$. Thus, the optimization problem reduces to

$$\min_{\mathbf{y}} \left\{ -2\mathbf{x}^\top \mathbf{y} \right\} \qquad \text{subject to} \qquad \|\mathbf{y}\|^2 = r^2\,.$$

Now, we will find a lower bound of the objective function using Cauchy–Schwarz,

$$-2\mathbf{y}^\top \mathbf{x} \geq -2\|\mathbf{y}\|\|\mathbf{x}\| = -2r\|\mathbf{x}\|$$

which is attained at $\mathbf{y} = -r\frac{\mathbf{x}}{\|\mathbf{x}\|}$. Summarizing,

$$\mathbb{P}_{B[\mathbf{0},r]}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \|\mathbf{x}\| \leq r, \\ r\dfrac{\mathbf{x}}{\|\mathbf{x}\|}, & \text{otherwise}\,. \end{cases}$$

and thus

$$\mathbb{P}_{B[\mathbf{x}_0,r]}(\mathbf{x}) = \begin{cases} \mathbf{x}, & \text{if } \|\mathbf{x}-\mathbf{x}_0\| \leq r, \\ \mathbf{x}_0 + r\dfrac{\mathbf{x}-\mathbf{x}_0}{\|\mathbf{x}-\mathbf{x}_0\|}, & \text{otherwise}\,. \end{cases}$$

2. Using the definition of stationarity of $\mathbf{x}^*$ over the unit ball, we have:

$$\nabla f(\mathbf{x}^*)^\top(\mathbf{x}-\mathbf{x}^*) \geq 0 \quad \forall \mathbf{x} \in B[0,1]\,.$$

This is equivalent to claim that

$$\min_{\mathbf{x}\in B[0,1]} \left\{ \nabla f(\mathbf{x}^*)^\top(\mathbf{x}-\mathbf{x}^*) \right\} \geq 0\,. \tag{1}$$

Lemma: *For any $\mathbf{a} \in \mathbb{R}^n$ we have*

$$\min_{\|\mathbf{x}\|\leq 1} \mathbf{a}^\top \mathbf{x} = -\|\mathbf{a}\|$$

which is attained at $\mathbf{x}^* = -\dfrac{\mathbf{a}}{\|\mathbf{a}\|}$. This can be shown as

$$\mathbf{a}^\top \mathbf{x} \geq -\|\mathbf{a}\|\|\mathbf{x}\| \geq -\|\mathbf{a}\|\,.$$

On one hand, the *Lemma* implies that (1) is equivalent to $-\nabla f(\mathbf{x}^*)^\top \mathbf{x}^* \geq \|\nabla f(\mathbf{x}^*)\|$. On the other hand, by Cauchy-Schwarz inequality, we have

$$-\nabla f(\mathbf{x}^*)^\top \mathbf{x}^* \leq \|\nabla f(\mathbf{x}^*)\|\|\mathbf{x}^*\| \leq \|\nabla f(\mathbf{x}^*)\|$$

as $\|\mathbf{x}^*\| \leq 1$. This leads to

$$-\nabla f(\mathbf{x}^*)^\top \mathbf{x}^* = \|\nabla f(\mathbf{x}^*)\|\,. \tag{2}$$

We now discuss two different cases:

a) $\nabla f(\mathbf{x}^*) = 0$ (and $\|\mathbf{x}^*\| \leq 1$) $\implies$ (2) holds;

b) $\nabla f(\mathbf{x}^*) \neq 0 \implies \|\mathbf{x}^*\| = 1$ and then

$$-\nabla f(\mathbf{x}^*)^\top \mathbf{x}^* = \|\nabla f(\mathbf{x}^*)\| \underbrace{\|\mathbf{x}^*\|}_{1} \iff \exists \lambda \leq 0 \;\text{ s.t. }\; \nabla f(\mathbf{x}^*) = \lambda \mathbf{x}^*$$

3. see `week8.m`

Extra exercise (7)

Given f, g convex functions over $\mathbb{R}^n$, $X \subseteq \mathbb{R}^n$ convex set, suppose $\mathbf{x}^*$ is a solution of

$$\min_{\mathbf{x} \in X} f(\mathbf{x}) \text{ subject to } \; g(\mathbf{x}) \leq 0 \tag{3}$$

that satisfies $g(\mathbf{x}^*) < 0$. Show that $\mathbf{x}^*$ is also a solution of

$$\min_{\mathbf{x} \in X} f(\mathbf{x})\,.$$

We assume there exists $\mathbf{y} \in X \cap \{\mathbf{x} : g(\mathbf{x}) > 0\}$ such that $f(\mathbf{y}) < f(\mathbf{x}^*)$. Both $\mathbf{x}^*, \mathbf{y} \in X$, which is convex, and $g(\mathbf{x}^*) < 0$ while $g(\mathbf{y}) > 0$. Due to continuity of g, we have that there exists a $\mathbf{z} \in [\mathbf{x}^*, \mathbf{y}] \in X$ such that $g(\mathbf{z}) = 0$, with $\mathbf{z} = \lambda \mathbf{y} + (1 - \lambda)\mathbf{x}^*$ for some $\lambda \in [0, 1]$. Since f is a convex function, we have

$$f(\mathbf{z}) = f(\mathbf{x}^* + \lambda(\mathbf{y} - \mathbf{x}^*)) \leq f(\mathbf{x}^*) + \underbrace{\lambda}_{>0} \overbrace{\big(f(\mathbf{y}) - f(\mathbf{x}^*)\big)}^{>0 \text{ by assumption}} < f(\mathbf{x}^*),$$

thus $f(\mathbf{z}) < f(\mathbf{x}^*)$ for a $z \in X$ such that $g(\mathbf{z}) = 0$. Since $\mathbf{z}$ belongs to the feasible set of (3), this leads to a contradiction to the optimality of $\mathbf{x}^*$ for (3).