DATA130026.01 Optimization
Solution of Assignment 9

1. Let f be a convex and continuous differentiable function over $\mathbb{R}^n$. For a fixed $x \in \mathbb{R}^n$, define the functions
$$g_x(y) = f(y) - \nabla f(x)^T y.$$
Suppose ∇f is L Lipschitz continuous, i.e.,
$$\|\nabla f(x) - \nabla f(y)\| \le L\|x - y\|, \quad \forall x, y \in \mathbb{R}^n.$$
   (a) Prove that x is a minimizer of g_x over $\mathbb{R}^n$.
   (b) Show that for any $x, y \in \mathbb{R}^n$,
   $$g_x(x) \le g_x(y) - \frac{1}{2L}\|\nabla g_x(y)\|^2.$$
   (c) Show that for any $x, y \in \mathbb{R}^n$,
   $$f(x) + \nabla f(x)^T(y - x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|^2 \le f(y).$$

Solution.

(a) Since $f(y)$ is convex and $\nabla f(x)^T y$ is affine, $g_x(y)$ is also convex. Let $\nabla g_x(y) = \nabla f(y) - \nabla f(x) = 0$, we have $y = x$. Thus, x is a minimizer of g_x over $\mathbb{R}^n$.

(b) $\forall\ y_1, y_2 \in \mathbb{R}^n$,
$$\begin{aligned}
\|\nabla g_x(y_1) - \nabla g_x(y_2)\| &= \|(\nabla f(y_1) - \nabla f(x)) - (\nabla f(y_2) - \nabla f(x))\| \\
&= \|\nabla f(y_1) - \nabla f(y_2)\| \\
&\le L\|y_1 - y_2\|.
\end{aligned}$$
Therefore, ∇g_x is also L Lipschitz continuous. Then, $\forall\ y, z \in \mathbb{R}^n$, we have
$$\begin{aligned}
& g_x(z) \le g_x(y) + \nabla g_x(y)^T(z - y) + \frac{L}{2}\|z - y\|^2, \\
\Rightarrow \quad & \min_z g_x(z) \le \min_z\{g_x(y) + \nabla g_x(y)^T(z - y) + \frac{L}{2}\|z - y\|^2\}, \\
\Rightarrow \quad & g_x(x) \le g_x(y) + \nabla g_x(y)^T\left(y - \frac{1}{L}\nabla g_x(y) - y\right) + \frac{L}{2}\left\|y - \frac{1}{L}\nabla g_x(y) - y\right\|^2 \\
& \quad = g_x(y) - \frac{1}{2L}\|\nabla g_x(y)\|^2.
\end{aligned}$$

(c) Since $g_x(x) = f(x) - \nabla f(x)^T x$ and $g_x(y) = f(y) - \nabla f(x)^T y$, from (b), we have
$$\begin{aligned}
& f(x) - \nabla f(x)^T x \le f(y) - \nabla f(x)^T y - \frac{1}{2L}\|\nabla g_x(y)\|^2, \\
\Rightarrow \quad & f(x) + \nabla f(x)^T(y - x) + \frac{1}{2L}\|\nabla f(x) - \nabla f(y)\|^2 \le f(y).
\end{aligned}$$

2. Let $F(x) = Ax + b$ be an affine function, with A an $n \times n$-matrix. What properties of the matrix A correspond to the following conditions (a)-(e) on F? Suppose that A is symmetric, so $F(x)$ is the gradient of a quadratic function.

   (a) Monotonicity:
   $$(F(x) - F(y))^T(x - y) \geq 0, \ \forall x, y.$$
   (b) Strict monotonicity:
   $$(F(x) - F(y))^T(x - y) > 0, \ \forall x, y.$$
   (c) Strong monotonicity (for the Euclidean norm):
   $$(F(x) - F(y))^T(x - y) \geq m\|x - y\|_2^2, \ \forall x, y,$$
   where m is a positive constant.

   (d) Lipschitz continuity (for the Euclidean norm):
   $$\|F(x) - F(y)\|_2 \leq L\|x - y\|_2, \ \forall x, y,$$
   where L is a positive constant.

   (e) Co-coercivity (for the Euclidean norm):
   $$(F(x) - F(y))^T(x - y) \geq \frac{1}{L}\|F(x) - F(y)\|_2^2, \ \forall x, y,$$
   where L is a positive constant.

   Solution.

   (a) For all x, y, we have
   $$(F(x) - F(y))^T(x - y) = (x - y)^T A(x - y) \geq 0,$$
   thus $A \succeq 0$.

   (b) For all x, y, we have
   $$(F(x) - F(y))^T(x - y) = (x - y)^T A(x - y) > 0,$$
   thus $A \succ 0$.

   (c) For all x, y, we have
   $$(F(x) - F(y))^T(x - y) - (x - y)^T mI(x - y) = (x - y)^T(A - mI)(x - y) \geq 0,$$
   thus $A - mI \succeq 0$. ($\lambda_{min}(A) \geq m$)

   (d) For all x, y, since $\|F(x) - F(y)\|_2^2 \leq L^2\|x - y\|_2^2$, we have
   $$(x - y)^T A^T A(x - y) - L^2(x - y)^T(x - y) = (x - y)^T(A^T A - L^2 I)(x - y) \leq 0,$$
   thus $L^2 I - A^T A \succeq 0$. ($\|A\|_2 \leq L$; $\lambda_{max}(A^2) \leq L^2$; $|\lambda|_{max}(A) \leq L$)

   (e) For all x, y, we have
   $$(F(x) - F(y))^T(x - y) - \frac{1}{L}\|F(x) - F(y)\|_2^2 = (x - y)^T\left(A - \frac{1}{L}A^T A\right)(x - y) \geq 0,$$
   thus $A - \frac{1}{L}A^T A \succeq 0$. ($\forall i, 0 \leq \lambda_i(A) \leq L$)