Lecture 2 Exercises

Exercise 2.1

Show that the update in (2.3) is indeed equivalent to the GD update.

Solution: Equation (2.3) is:

$$x_{k+1} = \underset{x}{\operatorname{argmin}} \left\{ f(x_k) + \langle \nabla f(x_k), x - x_k \rangle + \frac{1}{2\alpha_k} \|x - x_k\|^2 \right\}$$

To find the minimum, we take the derivative with respect to x and set it to zero:

$$\nabla f(x_k) + \frac{1}{\alpha_k}(x - x_k) = 0$$

$$x = x_k - \alpha_k \nabla f(x_k)$$

Thus, $x_{k+1} = x_k - \alpha_k \nabla f(x_k)$, which is the GD update.

Exercise 2.2

Show that $\alpha = \frac{1}{L}$ is indeed the maximum of $\alpha(2 - \alpha L)$.

Solution: We want to maximize $g(\alpha) = \alpha(2 - \alpha L) = 2\alpha - \alpha^2 L$. To find the maximum, we take the derivative with respect to α and set it to zero:

$$g'(\alpha) = 2 - 2\alpha L = 0$$

$$\alpha = \frac{1}{L}$$

To check that this is a maximum, we take the second derivative:

$$g''(\alpha) = -2L < 0$$

Since the second derivative is negative, $\alpha = \frac{1}{L}$ is indeed the maximum.

Exercise 2.3

Prove that f in (2.7) is convex if and only if $Q \succeq 0$.

Solution: Equation (2.7) is $f(x) = \frac{1}{2}\langle Qx, x \rangle - \langle b, x \rangle$. The Hessian of $f(x)$ is $\nabla^2 f(x) = Q$. A function is convex if and only if its Hessian is positive semi-definite, i.e., $Q \succeq 0$.

Exercise 2.4

Prove that f in (2.7) is L-smooth, with $L = ||Q|| = \lambda_{max}(Q)$.

Solution: We need to show that $||\nabla^2 f(x)|| \leq L$ for all x. We have $\nabla f(x) = Qx - b$, so $\nabla^2 f(x) = Q$. The L-smoothness condition is equivalent to $||\nabla^2 f(x)|| = ||Q|| \leq L$. The spectral norm of a symmetric matrix Q is its largest eigenvalue. Thus, $L = \lambda_{max}(Q)$.

Exercise 2.5

Explain all the steps in the proof of Lemma 2.3.

Solution: Lemma 2.3 states that if f is μ-strongly convex and x^* is the solution, then $\frac{1}{2\mu}||\nabla f(x)||^2 \geq f(x) - f^*$ for all x. The proof is as follows:

1. Start with the definition of μ-strong convexity:
$$f(x^*) \geq f(x) + \langle \nabla f(x), x^* - x \rangle + \frac{\mu}{2}||x - x^*||^2$$
2. Since x^* is the minimum, $\nabla f(x^*) = 0$.
3. Rearrange the inequality:
$$f(x) - f(x^*) \leq \langle \nabla f(x), x - x^* \rangle - \frac{\mu}{2}||x - x^*||^2$$
4. Use the inequality $\langle a, b \rangle \leq \frac{\epsilon}{2}||a||^2 + \frac{1}{2\epsilon}||b||^2$ with $a = x - x^*$ and $b = \nabla f(x)$, and $\epsilon = \frac{1}{\mu}$.
$$\langle \nabla f(x), x - x^* \rangle \leq \frac{1}{2\mu}||\nabla f(x)||^2 + \frac{\mu}{2}||x - x^*||^2$$
5. Substitute this back into the inequality:
$$f(x) - f(x^*) \leq \frac{1}{2\mu}||\nabla f(x)||^2 + \frac{\mu}{2}||x - x^*||^2 - \frac{\mu}{2}||x - x^*||^2$$
$$f(x) - f(x^*) \leq \frac{1}{2\mu}||\nabla f(x)||^2$$
6. Therefore, $\frac{1}{2\mu}||\nabla f(x)||^2 \geq f(x) - f^*$.

Exercise 2.6

Prove that $1 - t \leq e^{-t}$ using only convexity arguments (one line proof).

Solution: The function e^{-t} is convex on $\mathbb{R}$, and at $t = 0$, $e^{-0} = 1$, and the tangent line to e^{-t} at $t = 0$ is $1 - t$. Since the tangent line to a convex function lies below the function, we have $1 - t \leq e^{-t}$.