SOLUTIONS HW 6

1 Problem 1

For any $z \in \mathcal{Z}$, we have:

$$g(y,z) \le \max_{y\in\mathcal{Y}} g(y,z).$$

Next, we minimize both sides over $z \in \mathcal{Z}$, and the inequality still holds

$$\min_{z\in\mathcal{Z}} g(y,z) \le \min_{z\in\mathcal{Z}} \max_{y\in\mathcal{Y}} g(y,z).$$

The left side of the above inequality is a function of y, and the right side is a constant upper bound for the left side over all y. Therefore, the maximum of the left side over y should still be upper bounded by the constant on the right side. Hence we have

$$\max_{y\in\mathcal{Y}} \min_{z\in\mathcal{Z}} g(y,z) \le \min_{z\in\mathcal{Z}} \max_{y\in\mathcal{Y}} g(y,z).$$

2 Problem 2

The original problem is equivalent to

$$\begin{aligned} &\text{minimize } c^\mathsf{T} x \\ &\text{subject to } Ax - b \le 0, \end{aligned}$$

where $x = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, $c = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $A = \begin{bmatrix} -1 & -2 \\ 3 & 1 \\ -1 & 1 \end{bmatrix}$, and $b = \begin{bmatrix} -1 \\ 5 \\ 8 \end{bmatrix}$. To find the dual, for $\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix} \ge 0$, we calculate $D(\mu)$ as:

$$\begin{aligned} D(\mu) &= \min_{x\in\mathbb{R}^2} c^\mathsf{T} x + \mu^\mathsf{T}(Ax-b) \\ &= \min_{x\in\mathbb{R}^2} (c^\mathsf{T} + \mu^\mathsf{T} A)x - \mu^\mathsf{T} b \\ &= \begin{cases} -\infty & \text{if } c^\mathsf{T} + \mu^\mathsf{T} A \ne 0 \\ -\mu^\mathsf{T} b & \text{if } c^\mathsf{T} + \mu^\mathsf{T} A = 0. \end{cases} \end{aligned}$$

Notice that $c^\mathsf{T} + \mu^\mathsf{T} A = 0$ is equivalent to $A^\mathsf{T}\mu = -c$. Therefore, the dual problem is:

$$\begin{aligned} &\text{maximize } -\mu^\mathsf{T} b \\ &\text{subject to } A^\mathsf{T}\mu = -c,\ \mu \ge 0. \end{aligned}$$

To verify the strong duality, notice that the solution for the primal problem is given by $x = (-5, 3)$, and the optimal value for the primal problem is -2. For the dual problem, the optimal point is given by $\mu = (\frac{2}{3}, 0, \frac{1}{3})$, and the maximum value is $-\mu^\mathsf{T} b = -2$. Therefore, the primal and dual problems have the same solution. The strong duality holds.

Another way to verify strong duality is to use the Slater's condition. For linear programming, finding a strictly feasible point for the primal problem (e.g. $x = \begin{bmatrix} 1 & 1 \end{bmatrix}^\mathsf{T}$) does guarantee the strong duality to hold.

3 Problem 3

The Lagrangian of the problem is $L(x,\mu) = x^\intercal Q x + \mu^\intercal (Ax - b)$. Thus the Lagrangian dual function is:

$$\begin{aligned} D(\mu) &= \min_x L(x,\mu) \\ &= \min_x x^\intercal Q x + \mu^\intercal (Ax - b) \\ &= \min_x x^\intercal Q x + \mu^\intercal Ax - \mu^\intercal b. \end{aligned}$$

Since Q is positive definite, we can just take the derivative of $L(x,\mu)$ with respect to x and set it equal to 0. Then we obtain $x = -\frac{1}{2}Q^{-1}A^\intercal \mu$, which leads to:

$$D(\mu) = -\frac{1}{4}\mu^\intercal A Q^{-1} A^\intercal \mu - \mu^\intercal b.$$

Therefore, the dual problem is:

$$\text{maximize } \; -\frac{1}{4}\mu^\intercal A Q^{-1} A^\intercal \mu - \mu^\intercal b$$

4 Problem 4

Let $\bar{x}$ be a limit point of $\{x^{(k)}\}$ given by

$$\bar{x} = \min_{k\to\infty, k\in\mathcal{K}} x^{(k)}.$$

Assuming that $\min_{h(x)=0} f(x) = f^*$ exists, then we have:

$$\begin{aligned} f^* &= \min_{h(x)=0} f(x) \\ &= \min_{h(x)=0} f(x) + \lambda^\intercal h(x) + c_k \|h(x)\|^2 \\ &\geq \min f(x) + \lambda^\intercal h(x) + c_k \|h(x)\|^2 \\ &= f(x^{(k)}) + \lambda^\intercal h(x^{(k)}) + c_k \|h(x^{(k)})\|^2. \end{aligned}$$

This implies that

$$\begin{aligned} c_k \|h(x^{(k)})\|^2 + \lambda^\intercal h(x^{(k)}) &\leq f^* - f(x^{(k)}) \\ \Rightarrow c_k \|h(x^{(k)})\|^2 - \|\lambda\| \|h(x^{(k)})\| &\leq f^* - f(x^{(k)}) \\ \Rightarrow -\|\lambda\| \|h(x^{(k)})\| &\leq f^* - f(x^{(k)}), \end{aligned} \tag{1}$$

where the second step applies Cauchy–Schwarz inequality. By continuity of f, we have $\lim_{k\to\infty} f(x^{(k)}) = f(\bar{x})$. Thus as $k \to \infty$, $f^* - f(x^{(k)})$ goes to $f^* - f(\bar{x})$ which is finite. Since $c_k \to \infty$ as $k \to \infty$, we get

$$\lim_{k\to\infty, k\in\mathcal{K}} \|h(x^{(k)})\| = 0.$$

By continuity of $\|h(x)\|$, we get

$$\lim_{k\to\infty, k\in\mathcal{K}} \|h(x^{(k)})\| = \|h(\bar{x})\| = 0.$$

Taking limit as $k \to \infty, k \in \mathcal{K}$ in (1), we get

$$f^* - f(\bar{x}) \geq 0.$$

But $\bar{x}$ satisfies $h(\bar{x}) = 0$ and so $f(\bar{x}) \geq f^*$. Hence, every limit point is a global minimum.

5 Problem 5

(a)

$$
\begin{aligned}
& g \text{ is a subgradient of } f \text{ at } x \\
\iff & f(y) \geq f(x) + g^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
\iff & af(y) \geq af(x) + ag^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n, \;\; a > 0 \\
\iff & ag \text{ is a subgradient of } af \text{ at } x.
\end{aligned}
$$

(b) If g_1 is a subgradient of f_1 and g_2 is a subgradient of f_2 at x, then

$$
\begin{aligned}
& f_1(y) \geq f_1(x) + g_1^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
& f_2(y) \geq f_2(x) + g_2^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n \\
\Rightarrow \;\; & f_1(y) + f_2(y) \geq f_1(x) + f_2(x) + (g_1+g_2)^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n,
\end{aligned}
$$

which implies that $g_1 + g_2$ is a subgradient of $f_1 + f_2$ at x.

(c)

$$
\begin{aligned}
& g \text{ is a subgradient of } f \text{ at } Ax+b \\
\iff & f(y) \geq f(Ax+b) + g^\mathsf{T}(y-(Ax+b)), \;\; \forall\, y \in \mathbb{R}^n \\
\iff & f(Ay+b) \geq f(Ax+b) + g^\mathsf{T}(Ay+b-(Ax+b)), \;\; \forall\, y \in \mathbb{R}^n, \;\; \text{(since } A \text{ is invertible)} \\
\iff & h(y) \geq h(x) + (A^\mathsf{T} g)^\mathsf{T}(y-x), \;\; \forall\, y \in \mathbb{R}^n, \;\; \text{(here } h(y) = f(Ay+b)\text{)} \\
\iff & A^\mathsf{T} g \text{ is a subgradient of } h \text{ at } x.
\end{aligned}
$$

6 Problem 6

Inspired by the 1-D case, we can conjecture that any vector of the form $\begin{bmatrix} a & b & c \end{bmatrix}^\mathsf{T}$, with $a \in [-1, 1]$, $b \in [-1, 1]$, and $c \in [-1, 1]$ is a subgradient of f at the $(x_1, x_2, x_3) = (0, 0, 0)$. To this end, we need to show that for any $a \in [-1, 1]$, $b \in [-1, 1]$, and $c \in [-1, 1]$, we have:

$$
f(y) \geq f(0) + \begin{bmatrix} a & b & c \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix},
$$

which is equivalent as:

$$
|y_1| + |y_2| + |y_3| \geq ay_1 + by_2 + cy_3, \;\; \forall\, y \in \mathbb{R}^3. \tag{2}
$$

Since $a \in [-1, 1]$, $b \in [-1, 1]$, and $c \in [-1, 1]$, we have:

$$
ay_1 \leq |y_1|, \;\; by_2 \leq |y_2|, \;\; cy_3 \leq |y_3|.
$$

Hence (2) holds and $\begin{bmatrix} a & b & c \end{bmatrix}^\mathsf{T}$ is a subgradient of f at the origin. Finally, it is easy to show that if a, b or c are outside the interval $[-1, 1]$, then $\begin{bmatrix} a & b & c \end{bmatrix}^\mathsf{T}$ cannot be a subgradient of f at the origin.