MATH-GA.2012.001 Selected Topics in Numerical Analysis :
Convex and Nonsmooth Optimization, Spring 2020
Homework Assignment 3
Yves Greatti - yg390

1. (a) $L(x, \nu) = \frac{1}{2}x^T Qx + \nu^T(Ax - b)$

   (b) $\nabla_x L(x, \nu) = Qx + A^T\nu$, and $\nabla_x L(x, \nu) = 0 \Rightarrow x = -Q^{-1}A^T\nu$. For this minimizer x, the Lagrange dual function is: $g(\nu) = \min_x L(x, \nu) = \frac{1}{2}(-Q^{-1}A^T\nu)^T Q(-Q^{-1}A^T\nu) + \nu^T A(-Q^{-1}A^T\nu) - \nu^T b$

$$g(\nu) = -\frac{1}{2}\nu^T AQ^{-1}A^T\nu - \nu^T b$$

   (c)
$$\nabla_\nu g(\nu) = -AQ^{-1}A^T\nu - b$$
$$\nabla_\nu g(\nu) = 0 \Rightarrow \nu^* = -(AQ^{-1}A^T)^{-1}b$$

   The dual optimal value d^* is:

$$d^* = g(\nu^*) = -\frac{1}{2}b^T[-(AQ^{-1}A^T)^{-1}]^T AQ^{-1}A^T[-(AQ^{-1}A^T)^{-1}]b - b^T[-(AQ^{-1}A^T)^{-1}]^T b$$

$$d^* = \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}b$$

   (d) The associated $\hat{x}$ is:

$$\hat{x} = -Q^{-1}A^T\nu^* = -Q^{-1}A^T[-(AQ^{-1}A^T)^{-1}b] = Q^{-1}A^T(AQ^{-1}A^T)^{-1}b$$

   (e) $\hat{x}$ is feasible for the primal problem since it verifies the equality constraint $h(x) = Ax - b = 0$ since $h(\hat{x}) = AQ^{-1}A^T(AQ^{-1}A^T)^{-1}b - b = 0$.

   (f)

$$f_0(\hat{x}) = \frac{1}{2}\hat{x}^T\hat{x} = \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}AQ^{-1}QQ^{-1}A^T(AQ^{-1}A^T)^{-1}b$$

$$f_0(\hat{x}) = \frac{1}{2}b^T(AQ^{-1}A^T)^{-1}b = p^* = d^*$$

   (g) $d^* = p^*$ therefore there is no duality gap.

2. BV Ex 5.1

   (a) Given $f_0(x) = x^2 + 1$, the feasible set is the interval $[2, 4]$. $f_0'(x) = 2x$ so f_0 is increasing on $[2, 4]$, and the optimal value is $f_0(2) = 5$ and the optimal solution is 2.

(b)

$$L(x, \lambda) = x^2 + 1 + \lambda(x-1)(x-4) = (\lambda+1)x^2 - 6\lambda x + 1 + 8\lambda$$

The feasible set is indicated by the grey area between the two vertical lines at $x = 2$ and $x = 4$, $L(x, \lambda)$ is plotted for $\lambda = 1, 2, 3$ (ref **??**). We can verify on the graph that $\inf_x L(x, \lambda) \leq p^*, p^* = 5$. $L(x, \lambda)$ is convex (quadratic) and differentiable, we can set its gradient to zero to find its minimizer and write the Lagrange dual function $g(\nu)$. $\nabla_x L(x, \lambda) = 2(\lambda + 1)x - 6\lambda$, $\nabla_x L(x, \lambda) = 0 \Rightarrow x = \frac{3\lambda}{\lambda+1}$ for $\lambda \neq -1$. $g(\nu) = \inf_x L(x, \lambda)$ is not bounded when $\lambda \leq -1$ thus

$$g(\lambda) = \begin{cases} \dfrac{-9\lambda^2}{\lambda+1} + 1 + 8\lambda & \text{when } \lambda > -1 \\ -\infty & \text{otherwise} \end{cases}$$

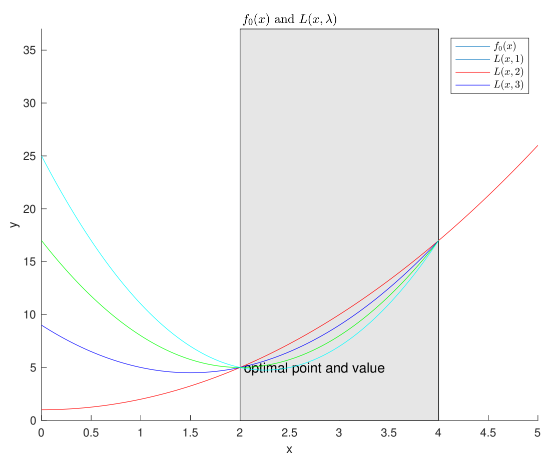

Figure 1: Plot of f_0 and the Lagragian $L(x, \lambda)$ for $\lambda = 1, 2, 3$.

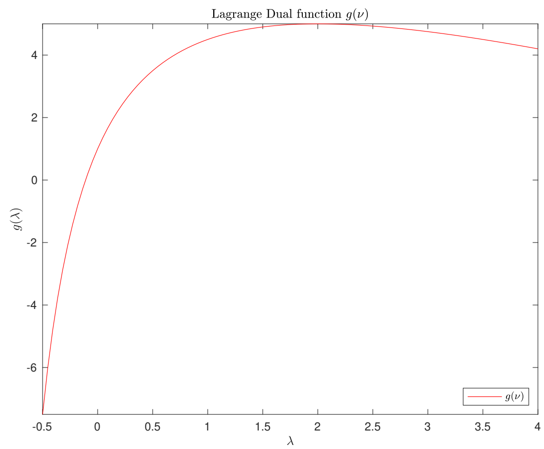

Figure 2: Plot of $g(\lambda)$ for $\lambda \in [-0.5, 5]$.

(c) The dual problem is:

$$\text{maximize } g(\lambda) = \frac{-9\lambda^2}{\lambda+1} + 1 + 8\lambda$$
$$\text{subject to } \lambda > -1$$

Based on the graph of $g(\lambda)$, it is concave. Setting the derivative of $g(\lambda)$ to zero, leads to $g'(\lambda) = 8 - 9\frac{\lambda(\lambda+2)}{(\lambda+1)^2} = 0$ which gives the dual optimal solution $\lambda^* = 2$ and the dual optimal value $d^* = 5 = p^*$ thus strong duality holds.

(d) Expanding $(x-2)(x-4) - u \leq 0$, we have $x^2 - 6x + 8 - u \leq 0$. The determinant $\Delta = 4(1+u)$ is valid only when $u \geq -1$ and the roots of the quadratic equation are $3 - \sqrt{1+u}$ and $3 + \sqrt{1+u}$, thus the feasible set is $[3-\sqrt{1+u}, 3+\sqrt{1+u}], u \geq -1$. We notice that $\min_x x^2 - 6x + 8 = -1$ and $p*(u)$ is not defined when $u < -1$. And for $p(u^*) = f(3-\sqrt{1+u}) = 11 + u - 6\sqrt{1+u}$ reaches its global minimizer point u_0^* when $p'(u_0^*) = 1 - \frac{3}{\sqrt{1+u}} = 0 \Rightarrow u_0^* = 8, p(u_0^*) = 1$. For $u \in [-1, 8]$, the solution to the problem in x is decreasing to the minimum value 1 to increase again when $u^* > 8$.

$$p(u^*) = \begin{cases} \infty & \text{when } u < -1 \\ 11 + u - 6\sqrt{1+u} & \text{when } u \in [-1, 8] \\ 1 & \text{when } u > 8 \end{cases}$$

We verify that $\frac{dp(u^*)}{du} = 1 - \frac{3}{\sqrt{1+u}}$ and $\frac{dp(0)}{du} = -2 = -\lambda^*$.

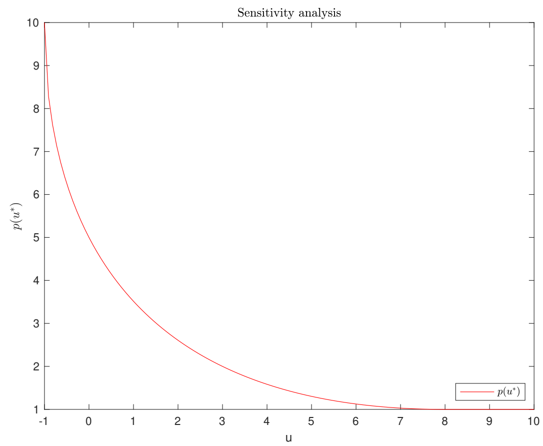

Figure 3: $p(u^*)$.

3. BV Ex 5.21

   (a) The exponential function e^{-x} is convex on $\mathbf{R}$, it then remains to determine if the Hessian of $f_1(x,y) = \frac{x^2}{y}$ is positive definite or semi definite.

   $$\nabla^2 f_1(x,y) = \begin{bmatrix} \frac{2}{y} & \frac{-2x}{y^2} \\ \frac{-2x}{y^2} & \frac{2x^2}{y^3} \end{bmatrix}$$

   The polynomial characteristic is $p(\lambda) = \frac{\lambda}{y^3}(y^3\lambda - 2(x^2+y^2))$ so the eigenvalues are 0 and $2\frac{x^2+y^2}{y^3} > 0, y > 0$. Hence the Hessian is positive semidefinite and $f_1(x,y)$ is convex. The problem is a convex optimization problem. The minimizer of e^{-x} satisfying the inequality constraint is $x = 0$ and the optimal value is 1.

   (b) The Lagrange dual function is $g(\lambda) = \inf_{x,y>0}(e^{-x} + \lambda\frac{x^2}{y})$

   $$g(\lambda) = \begin{cases} 0 & \text{when } \lambda \geq 0 \\ -\infty & \text{when } \lambda < 0 \end{cases}$$

   The optimal value is $d^* = 0$ for $\lambda^* \geq 0$, and the duality gap is $p^* - d^* = 1$.

   (c) Since there is not strong duality, Slater's condition cannot hold.

   (d) If $u < 0$ the problem is not defined then $p(u^*) = \infty$, if $u = 0, p(u^*) = 1$ and if $u > 0, p(u^*) = 0$. We observe that $p(u^*) \geq 1 - \lambda^* u$ since $\lambda^* \geq 0$.

4. Show that for convex problems of the form (5.25), that is (5.1) with $f_0, f_1, \ldots, f_m$ convex and $h_1, \ldots, h_p$ affine, the set $\mathcal{A} = \{(u,v,t) | \exists x \in \mathcal{D}, f_i(x) \leq u_i, i = 1, \ldots, m, h_i(x) = v_i, i = 1, \ldots, p, f_0(x) \leq t\}$ defined in BV (5.37) is convex.

Proof: Take two points

$$(u_1, v_1, t_1), (u_2, v_2, t_2) \in \mathcal{A}$$

$$\exists x_1, s.t. f_i(x_1) \leq u_{1_i}, i = 1, \ldots, m, h_i(x_1) = v_{1_i}, i = 1, \ldots, p, f_0(x_1) \leq t_1$$

and

$$\exists x_2, s.t. f_i(x_2) \leq u_{2_i}, i = 1, \ldots, m, h_i(x_2) = v_{2_i}, i = 1, \ldots, p, f_0(x_2) \leq t_2$$

$\forall \theta \in [0, 1]$, we have

$$\begin{aligned} f_i(\theta x_1 + (1-\theta)x_2) &\leq \theta f_i(x_1) + (1-\theta) f_i(x_2) \\ &\leq \theta u_{1_i} + (1-\theta) u_{2_i} \\ h_i(\theta x_1 + (1-\theta)x_2) &= \theta h_i(x_1) + (1-\theta) h_i(x_2) \\ &= \theta v_{1_i} + (1-\theta) v_{2_i} \\ f_0(\theta x_1 + (1-\theta)x_2) &\leq \theta f_0(x_1) + (1-\theta) f_0(x_2) \\ &\leq \theta t_1 + (1-\theta) t_2 \end{aligned}$$

$\theta(u_1, v_1, t_1) + (1-\theta)(u_2, v_2, t_2) \in \mathcal{A}$, thus $\mathcal{A}$ is convex.