Convex Optimization Quiz #4, Thursday June 6, 2019.

1. (48%) For each of the following optimization problems,

   (1) find dual function $g(\lambda, \nu)$ [1], and write down the dual problem (Each 8%).

   (2) write down the KKT conditions as up to five equations or inequalities[2] in x^*, λ^*, ν^* (Each 8%).

   (a) (16%)

$$\begin{array}{ll} \text{minimize} & \frac{1}{2}x^T P x + q^T x + r \\ \text{subject to} & Ax = b \end{array}$$

   (b) (16%)

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x \succeq 0 \end{array}$$

   where $A \in \mathbf{R}^{m \times n}$, $b \in \mathbf{R}^m$, and $c \in \mathbf{R}^n$.

   (c) (16%)

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, ..., n. \end{array}$$

   where $W \in \mathbf{S}^n$.

2. (40%) Consider the unconstrained problem we studied in class with the quadratic objective function on $\mathbf{R}^2$

$$f(x) = \frac{1}{2}(x_1^2 + \gamma x_2^2),$$

   where $\gamma > 0$.

   (a) (10%) Calculate the gradient $\nabla f(x)$ and the Hessian $\nabla^2 f(x)$.

   (b) (5%) Find the *condition number* of $\nabla^2 f(x)$.

   (c) (10%) Suppose we apply the *gradient descent method* with *exact line search* on this problem where the initial point is selected as $x^{(0)} = \begin{bmatrix} \gamma \\ 1 \end{bmatrix}$. That is, the $(k+1)$-st iterate can be expressed as

$$x^{(k+1)} = x^{(k)} + \left(\operatorname*{arg\,min}_{t>0} f(x^{(k)} + t\Delta x^{(k)}) \right) \cdot \Delta x^{(k)}$$

   where $\Delta x^{(k)} = -\nabla f(x^{(k)})$. Show that the kth iterate can be expressed in a closed form:

$$x^{(k)} = \begin{bmatrix} \gamma \left(\frac{\gamma-1}{\gamma+1}\right)^k \\ \left(-\frac{\gamma-1}{\gamma+1}\right)^k \end{bmatrix}.$$

   Hint: Find $t^{(k)} = \operatorname*{arg\,min}_t f(x^{(k)} + t \cdot \Delta x^{(k)})$ *first.*

[1] Note that $L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^m \lambda_i f_i(x) + \sum_{i=1}^p \nu_i h_i(x)$ and $g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu)$

[2] Hint: KKT conditions: (1)(2) primal feasibility; (3) dual feasibility; (4) complementary slackness; (5) the gradient of Lagrangian vanishing. The ith row of matrix A can be expressed as a_i^T.

(d) (10%) Calculate the Newton step $\Delta x_{\mathrm{nt}} = -(\nabla^2 f(x))^{-1}\nabla f(x)$ given any point $x \in \mathbf{R}^2$.

(e) (5%) How many iterations are needed for the Newton method with $t = 1$ to achieve an accurancy of $\epsilon = 10^{-6}$ on this problem? (Assume the initial point is selected as $x^{(0)} = \begin{bmatrix} \gamma \\ 1 \end{bmatrix}$.)

3. (12%) *True and False.* Determine the correctness of each of the following statements. No explanation needed. The score you'll get in this section is $\max\{0, 2n_c - 2n_w\}$ where n_c and n_w are the numbers of correct and wrong answers, respectively, not including those left blank. (i.e., $n_c + n_w \leq 6$).

   (a) (2%) The complementary slackness condition, $\sum_{i=1}^{m} \lambda_i^* f_i(x^*) = 0$, in a minimiztion problem, implies that "whenever $f_i(x^*) = 0$, then $\lambda_i^* > 0$."

   (b) (2%) The objective funcition in an uncontrained minimization problem is said to be *strongly convex* of S if there exists an $m > 0$ such that $\nabla^2 f(x) \succeq mI$ for all $x \in S$.

   (c) (2%) Gradient descent methods with backtracking line search usually converge faster, in terms of iteration numbers, than those with exact line search.

   (d) (2%) The dual norm of ℓ_p-norm, with $p > 1$, is the ℓ_q-norm where q satisfies $1/p + 1/q = 1$.

   (e) (2%) The steepest descent method coincides with the gradient descent method when the norm associated with the steepest descent method is chosen as the quadratic norm deteremined by the the Hessian of the objective function at the optimal point (i.e., $\nabla^2 f(x^*)$).

   (f) (2%) An important feature of the Newton step is that it is independent of linear (or affine) changes of coordinates.