CS5321 Numerical Optimization Homework 2

Due Dec 2

1. (20%) Check out the TRUST-REGION NEWTON- LANCZOS METHOD in Section 7.1 in the . What kind of problem it wants to solve? and how the Lanczos method solves it.

2. (20%) Check out section 8.2 in the deep learning textbook. Give a summary about the major challenges in neural network optimization.

3. (10%) Let J be an $m \times n$ matrix, $m \geq n$. Show that J has full column rank if and only if $J^T J$ is positive definite.

4. (30%) Simplex method (the algorithm is shown in Figure 2): Consider the following linear program:

$$
\begin{aligned}
\max_{x_1,x_2} \quad & z = 8x_1 + 5x_2 \\
\text{s.t.} \quad & 2x_1 + x_2 \leq 1000 \\
& 3x_1 + 4x_2 \leq 2400 \\
& x_1 + x_2 \leq 700 \\
& x_1 - x_2 \leq 350 \\
& x_1, x_2 \geq 0
\end{aligned}
$$

   (a) Transform it to the standard form.

   (b) Suppose the initial guess is $(0, 0)$. Use the simplex method to solve this problem. In each iteration, show

   - Basic variables and non-basic variables, and their values.
   - Pricing vector.
   - Search direction.
   - Ratio test result.

5. (20%) Farkas lemma: Let A be an $m \times n$ matrix and $\vec{b}$ be an m vector. Prove that exact one of the following two statements is true:

   (a) There exists a $\vec{x} \in \mathbb{R}^n$ such that $A\vec{x} = \vec{b}$ and $\vec{x} \geq 0$.

   (b) There exists a $\vec{y} \in \mathbb{R}^m$ such that $A^T \vec{y} \geq 0$ and $\vec{b}^T \vec{y} < 0$.

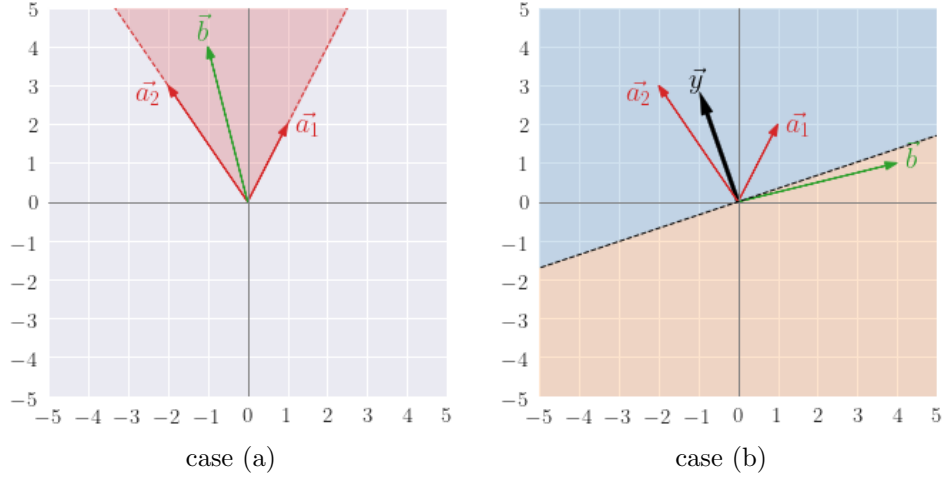   (Hint: prove if (a) is true, then (b) cannot be true, and vice versa.)

Figure 1: Two cases of Farkas Lemma.

Farkas's Lemma (1902) plays an important role in the proof of the KKT condition. The most critical part in the proof of the KKT condition is to show that the Lagrange multiplier $\vec{\lambda}^* \geq 0$ for inequality constraints. We can say if the LICQ condition is satisfied at $\vec{x}^*$, then any feasible direction $\vec{u}$ at $\vec{x}^*$ must have the following properties:

1. $\vec{u}^T \nabla f(\vec{x}^*) \geq 0$ since $\vec{x}^*$ is a local minimizer. (Otherwise, we find a feasible descent direction that decreases f.)

2. $\vec{u}^T \nabla c_i(\vec{x}^*) = 0$ for equality constraints, $c_i = 0$.

3. $\vec{u}^T \nabla c_i(\vec{x}^*) \geq 0$ for inequality constraints, $c_i \geq 0$.

Here is how Farkas Lemma enters the theme. Let $\vec{b}$ be $\nabla f(\vec{x}^*)$, $\vec{y}$ be $\vec{u}$ (any feasible direction at $\vec{x}^*$), the columns of A be $\nabla c_i(\vec{x}^*)$). Since no such $\vec{u}$ exists, according to the properties of $\vec{y}$, statement (a) must hold. The vector $\vec{x}$ in (a) corresponds to $\vec{\lambda}^*$, which just gives us the desired result of the KKT condition.

(1)	Given a basic feasible point $\vec{x}_0$ and the corresponding index set $\mathcal{B}_0$ and $\mathcal{N}_0$.
(2)	For $k = 0, 1, \ldots$
(3)	Let $B_k = A(:, \mathcal{B}_k), N_k = A(:, \mathcal{N}_k), \vec{x}_B = \vec{x}_k(\mathcal{B}_k), \vec{x}_N = \vec{x}_k(\mathcal{N}_k)$, and $\vec{c}_B = \vec{c}_k(\mathcal{B}_k), \vec{c}_N = \vec{c}_k(\mathcal{N}_k)$.
(4)	Compute $\vec{s}_k = \vec{c}_N - N_k^T B_k^{-1} \vec{c}_B$ (pricing)
(5)	If $\vec{s}_k \geq 0$, return the solution $\vec{x}_k$. (found optimal solution)
(6)	Select $q_k \in \mathcal{N}_k$ such that $\vec{s}_k(i_q) < 0$, where i_q is the index of q_k in $\mathcal{N}_k$
(7)	Compute $\vec{d}_k = B_k^{-1} A_k(:, q_k)$. (search direction)
(8)	If $\vec{d}_k \leq 0$, return `unbounded`. (unbounded case)
(9)	Compute $[\gamma_k, i_p] = \min\limits_{i, \vec{d}_k(i) > 0} \dfrac{\vec{x}_B(i)}{\vec{d}_k(i)}$ (ratio test) (The first return value is the minimum ratio; the second return value is the index of the minimum ratio.)
(10)	$x_{k+1}\begin{pmatrix} \mathcal{B} \\ \mathcal{N} \end{pmatrix} = \begin{pmatrix} \vec{x}_B \\ \vec{x}_N \end{pmatrix} + \gamma_k \begin{pmatrix} -\vec{d}_k \\ \vec{e}_{i_q} \end{pmatrix}$ ($\vec{e}_{i_q} = (0, \ldots, 1, \ldots, 0)^T$ is a unit vector with i_qth element 1.)
(11)	Let the i_pth element in $\mathcal{B}$ be p_k. (pivoting) $\mathcal{B}_{k+1} = (\mathcal{B}_k - \{p_k\}) \cup \{q_k\}, \mathcal{N}_{k+1} = (\mathcal{N}_k - \{q_k\}) \cup \{p_k\}$

Figure 2: The simplex method for solving (minimization) linear programming