PURDUE UNIVERSITY · CS 52000
COMPUTATIONAL OPTIMIZATION

HOMEWORK 5
Mathieu Gaillard
February 18, 2019

Problem 0: Homework checklist

I mainly used the class material, Google, Wikipedia and my previous knowledge.

Problem 1: Steepest descent

(Nocedal and Wright, Exercise 3.6) Let's conclude with a quick problem to show that steepest descent can converge very rapidly! Consider the steepest descent method with exact line search for the function $f(\mathbf{x}) = (1/2)\mathbf{x}^T\boldsymbol{Q}\mathbf{x} - \mathbf{x}^T\mathbf{b}$. Suppose that we know $\mathbf{x}_0 - \mathbf{x}^*$ is parallel to an eigenvector of $\boldsymbol{Q}$. Show that the method will converge in a single iteration.

$\mathbf{x}_0 - \mathbf{x}^*$ is parallel to an eigenvector of $\boldsymbol{Q}$ is equivalent to:

$$\exists\gamma \in \mathcal{R}, \lambda \in \mathcal{R}, \mathbf{y} \in \mathcal{R}^n, \quad \mathbf{x}_0 - \mathbf{x}^* = \gamma\mathbf{y} \quad \text{and} \quad \boldsymbol{Q}\mathbf{y} = \lambda\mathbf{y} \quad \text{and} \quad \|\mathbf{y}\| = 1$$

From the class material we know that:

$$\mathbf{x}_1 = \mathbf{x}_0 - \alpha\mathbf{g}_0$$
$$\mathbf{g}_0 = \boldsymbol{Q}\mathbf{x}_0 - \mathbf{b}$$
$$\alpha = \frac{\|\mathbf{g}_0\|^2}{\mathbf{g}_0^T\boldsymbol{Q}\mathbf{g}_0}$$

In $\mathbf{g}_0$, we replace $\mathbf{x}_0$ by $\gamma\mathbf{y} + \mathbf{x}^*$.

$$\begin{aligned}
\mathbf{g}_0 &= \boldsymbol{Q}\mathbf{x}_0 - \mathbf{b} \\
&= \boldsymbol{Q}(\gamma\mathbf{y} + \mathbf{x}^*) - \mathbf{b} \\
&= \gamma\boldsymbol{Q}\mathbf{y} + \boldsymbol{Q}\mathbf{x}^* - \mathbf{b} \\
&= \gamma\lambda\mathbf{y} + \boldsymbol{Q}\mathbf{x}^* - \mathbf{b} \\
&= \gamma\lambda\mathbf{y}
\end{aligned}$$

In the previous set of equations, $\mathbf{g}^* = \boldsymbol{Q}\mathbf{x}^* - \mathbf{b} = 0$ because according to the necessary conditions, the gradient at the optimum is equal to zero.

We substitute the value of α and $\mathbf{g}_0$ in $\mathbf{x}_1 = \mathbf{x}_0 - \alpha\mathbf{g}_0$.

$$
\begin{aligned}
\mathbf{x}_1 &= \mathbf{x}_0 - \alpha \mathbf{g}_0 \\
&= \mathbf{x}_0 - \frac{\gamma^2\lambda^2}{\gamma\lambda\mathbf{y}^T \boldsymbol{Q} \gamma\lambda\mathbf{y}} \times \gamma\lambda\mathbf{y} \\
&= \mathbf{x}_0 - \frac{\gamma^2\lambda^2}{\gamma^2\lambda^3} \times \gamma\lambda\mathbf{y} \\
&= \mathbf{x}_0 - \gamma\mathbf{y} \\
&= \mathbf{x}^*
\end{aligned}
$$

Therefore, the algorithm converges in only one iteration.

Problem 2: LPs in Standard Form

Show that we can solve:

$$
\begin{aligned}
\text{minimize} \quad & \|\mathbf{x}\|_1 + \|\mathbf{x}\|_\infty \\
\text{subject to} \quad & \boldsymbol{A}\mathbf{x} = \mathbf{b}
\end{aligned}
$$

by constructing an LP in standard form.

We convert the infinite norm into bounds:

$$
\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{x}\|_\infty \iff \begin{aligned} \underset{t}{\text{minimize}} \quad & t \\ \text{subject to} \quad & \|\mathbf{x}\|_\infty \leq t \end{aligned}
$$

$$
\begin{aligned}
\|\mathbf{x}\|_\infty \leq t &\iff \max_i \|x_i\| \leq t \\
&\iff \forall i \in 1, \ldots, n, -t \leq x_i \leq t \\
&\iff t\mathbf{e} - \mathbf{x} \geq 0, \quad t\mathbf{e} + \mathbf{x} \geq 0, \quad t \geq 0
\end{aligned}
$$

We convert the L1 norm into bounds:

$$
\underset{\mathbf{x}}{\text{minimize}} \quad \|\mathbf{x}\|_1 \iff \begin{aligned} \underset{\mathbf{y}}{\text{minimize}} \quad & \mathbf{e}^T\mathbf{y} \\ \text{subject to} \quad & -\mathbf{y} \leq \mathbf{x} \leq \mathbf{y}, \mathbf{y} \geq 0 \end{aligned}
$$

Our initial problem becomes:

$$
\begin{aligned}
\text{minimize} \quad & \mathbf{e}^T\mathbf{y} + t \\
\text{subject to} \quad & \boldsymbol{A}\mathbf{x} = \mathbf{b} \\
& \mathbf{y} - \mathbf{x} \geq 0 \\
& \mathbf{y} + \mathbf{x} \geq 0 \\
& \mathbf{y} \geq 0 \\
& t\mathbf{e} - \mathbf{x} \geq 0 \\
& t\mathbf{e} + \mathbf{x} \geq 0 \\
& t \geq 0
\end{aligned}
$$

$\mathbf{x}$ is unconstrained in sign, we convert it into a pair of positive variables:

$$
\exists \mathbf{x}^+ \geq 0, \mathbf{x}^- \geq 0 \quad \text{s.t.} \quad \mathbf{x} = \mathbf{x}^+ - \mathbf{x}^-
$$

We also add slack variables, the constraints become:

$$\boldsymbol{A}\mathbf{x}^+ - \boldsymbol{A}\mathbf{x}^- = \mathbf{b}$$

$$\mathbf{y} - \mathbf{x}^+ + \mathbf{x}^- - \mathbf{s}_1 = 0$$

$$\mathbf{y} + \mathbf{x}^+ - \mathbf{x}^- - \mathbf{s}_2 = 0$$

$$t\mathbf{e} - \mathbf{x}^+ + \mathbf{x}^- - \mathbf{s}_3 = 0$$

$$t\mathbf{e} + \mathbf{x}^+ - \mathbf{x}^- - \mathbf{s}_4 = 0$$

$$\mathbf{x}^+, \mathbf{x}^-, \mathbf{y}, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4, t \geq 0$$

Expressed as matrices:

$$\hat{\boldsymbol{A}} = \begin{bmatrix} \boldsymbol{A} & -\boldsymbol{A} & 0 & 0 & 0 & 0 & 0 & 0 \\ -I & I & I & 0 & -I & 0 & 0 & 0 \\ I & -I & I & 0 & 0 & -I & 0 & 0 \\ -I & I & 0 & \mathbf{e} & 0 & 0 & -I & 0 \\ I & -I & 0 & \mathbf{e} & 0 & 0 & 0 & -I \end{bmatrix}$$

$$\hat{\mathbf{x}} = (\mathbf{x}^+, \mathbf{x}^-, \mathbf{y}, t, \mathbf{s}_1, \mathbf{s}_2, \mathbf{s}_3, \mathbf{s}_4)^T$$

$$\hat{\mathbf{b}} = (\mathbf{b}, 0, 0, 0, 0)^T$$

$$\hat{\mathbf{c}} = (0, 0, \mathbf{e}, 1, 0, 0, 0, 0)^T$$

Now the standard form of the initial LP problem is:

$$\begin{array}{ll} \underset{\hat{\mathbf{x}}}{\text{minimize}} & \hat{\mathbf{c}}^T\hat{\mathbf{x}} \\ \text{subject to} & \hat{\boldsymbol{A}}\hat{\mathbf{x}} = \hat{\mathbf{b}} \\ & \hat{\mathbf{x}} \geq 0 \end{array}$$

Problem 3: Duality

Show that the these two problems are dual by showing the equivalence of the KKT conditions:

$$\begin{array}{ll} \underset{\mathbf{x}}{\text{minimize}} & \mathbf{c}^T\mathbf{x} \\ \text{subject to} & \boldsymbol{A}\mathbf{x} \geq \mathbf{b}, \mathbf{x} \geq 0 \end{array} \quad \text{and} \quad \begin{array}{ll} \underset{\boldsymbol{\lambda}}{\text{maximize}} & \mathbf{b}^T\boldsymbol{\lambda} \\ \text{subject to} & \boldsymbol{A}^T\boldsymbol{\lambda} \leq \mathbf{c}, \boldsymbol{\lambda} \geq 0 \end{array}.$$

We know from the class material that the KKT conditions for the primal problem are:

$$\mathbf{c} = \boldsymbol{A}^T\lambda + \mathbf{s}$$

$$\boldsymbol{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{x}^T\mathbf{s} = 0$$

$$\mathbf{s} \geq 0$$

$$\mathbf{x} \geq 0$$

The Lagrangian for the dual problem is:

$$\mathcal{L} = -\mathbf{b}^T\lambda - \mathbf{x}^T(\mathbf{c} - \boldsymbol{A}^T\lambda)$$

Then the KKT conditions for the dual problem are:

$$\frac{\partial \mathcal{L}}{\partial \lambda} = -\mathbf{b} + \boldsymbol{A}\mathbf{x} = 0$$

$$\boldsymbol{A}^T \lambda \leq \mathbf{c}$$

$$\mathbf{x} \geq 0$$

$$\mathbf{x}^T(\mathbf{c} - \boldsymbol{A}^T \lambda) = 0$$

As written in the class, if we substitute the variables as: $\mathbf{s} = \mathbf{c} - \boldsymbol{A}^T \lambda$, we get:

$$\mathbf{s} = \mathbf{c} - \boldsymbol{A}^T \lambda$$

$$\boldsymbol{A}\mathbf{x} = \mathbf{b}$$

$$\mathbf{s} \geq 0$$

$$\mathbf{x} \geq 0$$

$$\mathbf{x}^T \mathbf{s} = 0$$

These conditions are equivalent to the KKT conditions of the primal problem.

Problem 4: Geometry of LPs

(Griva, Sofer, and Nash, Problem 3.12) Consider the system of constraints $\boldsymbol{A}\mathbf{x} = \mathbf{b}, \mathbf{x} \geq 0$ with

$$\boldsymbol{A} = \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix}, \text{ and } \mathbf{b} = \begin{bmatrix} 12 \\ 15 \\ 18 \end{bmatrix}$$

Is $\mathbf{x} = \begin{bmatrix} 1 & 1 & 1 & 0 & 0 & 0 \end{bmatrix}^T$ a basic feasible point? Explain your answer precisely in terms of the definition.

First, we check that $\mathbf{x}$ is a feasible point. $\boldsymbol{A}\mathbf{x} = \mathbf{b}$ and $\mathbf{x} \geq 0$ so, yes it is a feasible point. With the script given in class, we enumerate all the basic feasible points and $\mathbf{x}$ is not among them. Why? Because the corresponding basis $\boldsymbol{B}$ would be the first three vectors of $\boldsymbol{A}$.

$$\boldsymbol{B} = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

However, $\boldsymbol{B}$ is singular because $\det \boldsymbol{B} = 0$. Thus, by definition it cannot be a basic feasible point.

Problem 5: Using the geometry

(Griva, Sofer, and Nash, Section 4.3, problem 3.13. Suppose that a linear program originally included a free variable x_i where there were no upper-and-lower bounds on its values. As we described in class, this can be converted into a pair of variables x_i^+ and x_i^- such that $x_i^+, x_i^- \geq 0$ and x_i is replaced with the difference $x_i^+ - x_i^-$. Prove that a basic feasible point can have only one of x_i^+ or x_i^- different from zero. (Hint: this is basically a one-line proof once you see the right characterization. I would suggest trying an example.)

Because $\mathbf{x} = \mathbf{x}^+ - \mathbf{x}-$, when we convert the problem into standard form, in every equation including $\mathbf{x}$ there are $\mathbf{x}^+$ and $\mathbf{x}-$ terms. Thus, in the updated $\boldsymbol{A}$ matrix, the two columns corresponding respectively to $\mathbf{x}^+$ and $\mathbf{x}-$ are going to be opposite vectors (linearly dependent). That's why a basic feasible point can have only one of them, because otherwise, the matrix $\boldsymbol{B}$ would be singular.