HOMEWORK 5
Elena Gómez
February 13, 2023

PROBLEM 0: Homework checklist
I worked alone.

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 $f(x) = (1/2)x^T Qx - x^T b$. Suppose that we know $x_0 - x^*$ is parallel to an eigenvector of Q. Show that the method will converge in a single iteration.

If we know that $x_0 - x^*$ is parallel to an eigenvector of Q, this means that
$x_0 - x^* = \mu v$
Where
$v \in R^n$ is an eigenvector of Q
$\mu \in R$ is a scalar
As v is an eigenvector of Q
$Qv = \lambda v$
where λ is the eigenvalue of Q for the eigenvector v

After explaining the supposition, let's show that the steepest descendent can converge in a single operation

Steepest descendent begins with some prescripted point x_0. At each step, it considers a linear approximation $f(x)$ in the direction of the negative gradient.
$x_{k+1} = x_k - \alpha_k g_k$.
So, the first iteration is computing as follows:
$x_1 = x_0 - \alpha_0 g_0$

Let's compute α_0 and g_0 in order to compute x_1
Firstly, let's compute g_0
$g_0 = Qx_0 - b$
as we know that $x_0 - x^* = \mu v$,we can say that $x_0 = \mu v - x^*$
So, let's substitubte x_0 by $\mu v - x^*$:
$g_0 = Q(\mu v - x^*) - b = \mu Qv + Qx^* - b$
As $Qv = \lambda v$:
$g_0 = \mu\lambda v + Qx^* - b$
Note that $Qx^* - b$ is the gradient at the optimum, $g^* = Qx^* - b$. The gradient as the optimum is 0, so $g^* = Qx^* - b = 0$ and:
$g_0 = \mu\lambda v$
Now, let's compute α_0
We know from class that
$\alpha = \frac{||g_k||^2}{g_k^T Q g_k}$,
So:
$\alpha_0 = \frac{||g_0||^2}{g_0^T Q g_0} = \frac{\mu^2\lambda^2 v^2}{(\mu\lambda v)^T Q \mu\lambda v} = \frac{1}{\lambda}$
because $Qv = \lambda v$

Now, let's compute x_1
$x_1 = x_0 - \alpha_0 g_0 = x_0 - \frac{1}{\lambda} \cdot \mu\lambda v = x_0 - \mu v$
As $x_0 - x^* = \mu v$, $x^* = x_0 - \mu v$
So :
$x_1 = x_0 - \mu v = x^*$

So, the method converges in a single iterarion.

PROBLEM 2: LPs in Standard Form
Show that we can solve:
$minimize \sum_i max(a_i^T x - b_i, -0.5)$
by constructing an LP standard form

The standard form for a linear program is

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^t x \\ \text{subject to} \quad & Ax = b \; and \; x \geq 0 \end{aligned}$$

In the previous hw (hw 4), we reformulated this as a constrained optimization problem:
Notice that: $\forall i \in \{1, 2, 3, \ldots, n\}$, $a_i^T x - b_i \geq -0.5 \rightarrow max(a_i^T x - b_i, -0.5) = a_i^T x - b_i$

$$\begin{aligned} \underset{z}{\text{minimize}} \quad & e^t z \\ \text{subject to} \quad & z = a_i^T x - b_i \; and \; z \geq -0.5. \end{aligned}$$

Let's define $\hat{c}$,$\hat{x}$,$\hat{A}$ and $\hat{b}$ in order to write our problem in LP standard form

$\hat{c} = (0, 0, e, -e, 0)^T$
$\hat{x} = (x^+, x^-, z^+, z^-, s)$
$\hat{b} = (b, 0.5)$

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

So, we get from
$\hat{A}\hat{x} = \hat{b}$:
2 equalities
The first one is:
$Ax^+ - Ax^- - z^+ + z^- = b$

because we want $z = a_i^T x - b$, so $-z + a_i^T x - b = 0$ and as x is uncontrained in sign
$-z + a_i^T x^+ - a_i^T x^- - b = 0$. We include z^+ and z^- because z can be a value in [-0.5, + $\infty)$

The second one is:
$z^- + s = 0.5$

because we want $z \geq -0.5$, so $-z^- + 0.5 - s = 0$

And from $c^T x$, we get:
$c^T x = e^T z^+ - e^t z^-$

As we will say that $\hat{x} \geq 0$, $x^+, x^-, z^+, z^-, s \geq 0$

So, the problem can be written in LP standard form as follows:

$$\begin{array}{ll} \underset{\hat{x}}{\text{minimize}} & \hat{c}^t \hat{x} \\ \text{subject to} & \hat{A}\hat{x} = \hat{b} \; and \; \hat{x} \geq 0 \end{array}$$

Thus, we have shown that we can solve the given problem using an LP standard form.

PROBLEM 3: Duality

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

$$\begin{array}{ll} \underset{x}{\text{minimize}} & c^t x \\ \text{subject to} & Ax \geq b, x \geq 0 \end{array}$$

and

$$\begin{array}{ll} \underset{\lambda}{\text{maximize}} & b^T \lambda \\ \text{subject to} & A^T \lambda \leq c, \lambda \geq 0 \end{array}$$

Let's see the KKT conditions of both problem in order to see if they are equivalent

KKT conditions for problem 1
Problem 1 is the standard form for a linear programming and we know from class that the KKt conditions are the following ones:
$A^T \lambda + s = c$
$Ax = b$

$x \geq 0$
$s \geq 0$
$x^T s = 0$

KKT conditions for problem 2

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & b^T \lambda \\ \text{subject to} \quad & A^T \lambda \leq c, \lambda \geq 0 \end{aligned}$$

is the same as

$$\begin{aligned} \underset{\lambda}{\text{minimize}} \quad & -b^T \lambda \\ \text{subject to} \quad & A^T \lambda \leq c, \lambda \geq 0 \end{aligned}$$

So, The Lagrangian is
$L = -b^T \lambda - x^T (c - A^T \lambda)$
And
$\frac{\partial L}{\partial \lambda} = -b + Ax = 0$

So, the KKT conditions are:
$Ax = b$
$A^T \lambda \leq c$
$x \geq 0$
$x^T (c - A^T \lambda) = 0$
And if we include the condition $c - A^T \lambda = s$,the KKT conditions are the following ones:
$A^T \lambda + s = c$
$Ax = b$
$x \geq 0$
$s \geq 0$
$x^T s = 0$

So, the KKT conditions of both problems are equivalent

PROBLEM 4: Geometry of LPs
(Griva, Sofer, and Nash, Problem 3.12) Consider the system of constraints $Ax = b, x \geq 0$ with
$A = \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 0 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix}$**, and** $b = \begin{bmatrix} 12 \\ 15 \\ 18 \end{bmatrix}$

Is $x = [1 \;\; 1 \;\; 1 \;\; 0 \;\; 0 \;\; 0]^T$ a basic feasible point? Explain your answer precisely in terms of the definition.

Fistly, let's see if x is a feasible point.
In order to be a feasible point the following conditions have to be satisfied:
$x \geq 0$
$Ax = b$
First condition: $x \geq 0$
It is satisfied since all entries of x are non-negative\

Second condition: $Ax = b$

$$Ax = \begin{bmatrix} 1 & 4 & 7 & 1 & 0 & 0 \\ 2 & 5 & 8 & 0 & 1 & 1 \\ 3 & 6 & 9 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ 1 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 12 \\ 15 \\ 18 \end{bmatrix} = b$$

So, the second condition is also satisfied and x is a feasible point

Now, let's see if x is a basic feasible point.
The indices corresponding to the non-zero entries of x are 1,2 and 3.
We have to see if the columns of A corresponding to indices 1,2 and 3 are linearly independent in order to see if this set forms a basis for A
In order to see if the columns corresponding to indices 1,2 and 3 are linearly independent, we create a new matrix B whose colums are the 3 first column of A and we compute its determinant.

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

Its determinant is:
$|B| = 0$
As $|B| = 0$, the first 3 columns of A are linearly dependent and indices 1,2 and 3 does not form a basis for A.
So, $x = [1 \ \ 1 \ \ 1 \ \ 0 \ \ 0 \ \ 0]^T$ is not a basic feasible point in this case

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^- ≥ 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.)

Suppose we have a basic feasible point with both x^+ and x^- non-zero. We express $x = x^+ - x^-$.
A basic feasible point is a vertex of the feasible region, where at least n linearly independent constraints are active. In the case of linear programming problems in standard form, the constraints are of the form $Ax = b$, $x \geq 0$, where A is an $m \times n$ matrix and b is a vector of length m.
If we suppose that there is a basic feasible point x with both x^+ and x^- non-zero, the columns of A corresponding to x^+ and x^- are opposite vectors,so the rows of B are linearly dependent, and B is non-invertible.
Therefore, a basic feasible point can have only one of x^+ or x^- different from zero, because if not B would be non-invertible.