HOMEWORK
David F. Gleich
February 13, 2023

Please answer the following questions in complete sentences in a clearly prepared manuscript and submit the solution by the due date on Gradescope.

Remember that this is a graduate class. There may be elements of the problem statements that require you to fill in appropriate assumptions. You are also responsible for determining what evidence to include. An answer alone is rarely sufficient, but neither is an overly verbose description required. Use your judgement to focus your discussion on the most interesting pieces. The answer to "should I include 'something' in my solution?" will almost always be: Yes, if you think it helps support your answer.

Problem 0: Homework checklist

- Please identify anyone, whether or not they are in the class, with whom you discussed your homework. This problem is worth 1 point, but on a multiplicative scale.
- Make sure you have included your source-code and prepared your solution according to the most recent Piazza note on homework submissions.

Problem 1: Log-barrier terms

The basis of a class of methods known as interior point methods is that we can handle non-negativity constraints such as $\mathbf{x} \geq 0$ by solving a sequence of unconstrained problems where we add the function $b(\mathbf{x}; \mu) = -\mu \sum_i \log(x_i)$ to the objective. Thus, we convert

$$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \geq 0 \end{array}$$

into

$$\text{minimize} \quad f(\mathbf{x}) + b(\mathbf{x}; \mu).$$

1. Explain why this idea could work. (Hint: there's a very useful picture you should probably show here!)

   Let's assume that $\mu > 0$ is small, then by minimizing $f(x) + b(x; \mu)$ we would obtain the same result as the original problem as the impact of the value X could be ignored. Secondly, for X on the left plane that fail the constraint $x \geq 0$, $b(x; \mu)$ is undefined which guarantees the points to dwell in the constraint region. Thus as long as μ is small, we can solve the original contrained problem by minimizing the new unconstrained: $f(\mathbf{x}) + b(\mathbf{x}; \mu)$. The parameter μ allows to control how much inside the feasible set we want to be, with respect to how much we insist on minimizing the objective; interior-point algorithms work by solving a sequence of unconstrained problems with increasing values of μ.

2. Write a matrix expression for the gradient and Hessian of $f(\mathbf{x}) + b(\mathbf{x}; \mu)$ in terms of the gradient vector $g(\mathbf{x})$ and the Hessian matrix $\boldsymbol{H}(\mathbf{x})$ of f.

   Let assume that$\mathbf{x} = (x_1, ..., xn)^T \in R_n$, and $g(x) = (g_1, ..., g_n)^T$ and $Q(x) = f(x) + b(x; \mu)$. thus the gradient of the unconstrained function $Q(x)$ can be formulated as:

$$\nabla Q(x) = g(x) + \partial b(x; \mu) = g(x) + \mu(\frac{1}{x_1}, ..., \frac{1}{x_n}) \tag{1}$$

and the Hessian of the unconstrained function $Q(x)$ can be formulated as:

$$\nabla H_Q(x) = H(x) + Hessian(b(x;\mu)) = H(x) - \mu diag(\frac{1}{x_1^2}, ..., \frac{1}{x_n^2}) \tag{2}$$

3. (Open ended) Explain if there is anything special about log that makes it especially useful here. For instance, are there other functions that could work and give you the same effect? Would log be superior?

   Since to convert the original problem into an unconstrained one, the barrier should go to infinity as $x \to \infty$ and it should go to zero as $x \to 0$. Since the logarithmic functions conform to these requirements, they are especially useful here. Any one candidate from the family of logarithmic functions would be a substitute here.

Problem 2: Inequality constraints

Draw a picture of the feasible region for the constraints:

$$\begin{bmatrix} 1 - x_1 - x_2 \\ 1 - x_1 + x_2 \\ 1 + x_1 - x_2 \\ 1 + x_1 + x_2 \end{bmatrix} \geq 0.$$

Problem 3: Necessary and sufficient conditions

Let $f(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T \boldsymbol{Q}\mathbf{x} - \mathbf{x}^T\mathbf{c}$.

1. Write down the necessary conditions for the problem:

   $$\begin{array}{ll} \text{minimize} & f(\mathbf{x}) \\ \text{subject to} & \mathbf{x} \geq 0 \end{array}.$$

   Let $g(x)$ denote the gradient of $f(x)$, thus the necessary condition for the problem is that there exists a Lagrangian multiplier λ that satisfies $\lambda = g(x), \lambda^T x = 0, \lambda \geq 0, x \geq 0$

2. Write down the sufficient conditions for the same problem.

   Let $H(x)$ denote the Hessian matrix of $f(x)$, the sufficient condition for the problem is that there exists a Lagrangian multiplier λ that satisfies the necessary condition and for any vector v in the critical cone, we have $v^T H(x) v$.

3. Consider the two-dimensional case with

   $$\boldsymbol{Q} = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix} \qquad \mathbf{c} = \begin{bmatrix} -1.5 \\ 1.5 \end{bmatrix}.$$

   Determine the solution to this problem by any means you can, and justify your work.

4. Produce a Julia or hand illustration of the solution showing the function contours, and gradient. What are the active constraints at the solution? What is the value of λ in $\boldsymbol{A}^T\lambda = \mathbf{g}$?

5. What changes when we set $\boldsymbol{Q} = \begin{bmatrix} -2 & 3 \\ 3 & -2 \end{bmatrix}$?

Problem 4: Constraints can make a non-smooth problem smooth.

Show that

$$\text{minimize} \quad \sum_i \max(\mathbf{a}_i^T \mathbf{x} - b_i, -0.5)$$

can be reformulated as a constrained optimization problem with a continuously differentiable objective function and both linear equality and inequality constraints.