Numerical analysis: Homework 5 solutions

Instructor: Anil Damle
Due: April 26, 2024

Policies

You may discuss the homework problems freely with other students, but please refrain from looking at their code or writeups (or sharing your own). Ultimately, you must implement your own code and write up your own solution to be turned in. Your solution, including plots and requested output from your code should be typeset and submitted via the Gradescope as a pdf file. This file must be self contained for grading. Additionally, please submit any code written for the assignment as zip file to the separate Gradescope assignment for code.

Question 1:

Consider the penalized formulation for solving a quadratic problem with equality constraints

$$\min_x \frac{1}{2}x^THx + x^Tc + \frac{1}{2\mu}\|Ax-b\|_2^2, \tag{1}$$

where $H \in \mathbb{R}^{n\times n}$ is symmetric, $c \in \mathbb{R}^n$, $\mu > 0$, and $A \in \mathbb{R}^{m\times n}$ with $m < n$ and full row-rank. Let $Z \in \mathbb{R}^{n\times(n-m)}$ be a matrix with orthonormal columns whose range is the null space of A. Show that if Z^THZ is positive definite, then for all sufficiently small μ (1) has a unique solution and that solution is a strict local minimizer.

Solution: Because the objective function is quadratic, it suffices to show that the Hessian matrix (which is constant in x) is positive definite. first, observe that the Hessian is

$$H + \frac{1}{2\mu}A^TA.$$

Now, we will directly show that this matrix is positive definite, i.e, that

$$x^T(H + \frac{1}{2\mu}A^TA)x > 0$$

for any $x \neq 0$. In fact, it suffices to show this for all x with $\|x\|_2 = 1$ since the sign is invariant to scaling of x. Given the assumption that Z^THZ is positive definite, it is reasonable to split x into the part in the null space of A and and the part perpendicular to that, denoted x_n and x_p respectively. Recall that for any x we have that x_n and x_p are unique and that $\|x\|_2^2 = \|x_n\|_2^2 + \|x_p\|_2^2$.

Now observe that for any x we have that

$$\begin{aligned}
x^T(H + \frac{1}{2\mu}A^TA)x &= (x_n + x_p)^T(H + \frac{1}{2\mu}A^TA)(x_n + x_p) \\
&= x_n^THx_n + 2x_n^THx_p + \frac{1}{2\mu}x_p^TA^TAx_p \\
&\geq \lambda_{\min}(Z^THZ)\|x_n\|_2^2 - 2\sigma_{\max}(H)\|x_p\|_2\|x_n\|_2 + \frac{1}{2\mu}\sigma_{\min}(A)\|x_p\|_2^2,
\end{aligned}$$

where by assumption $\lambda_{\min}(Z^T HZ) > 0$. Using the fact that $1 = \|x_n\|_2^2 + \|x_p\|_2^2$. we have that

$$\begin{aligned}
x^T(H + \frac{1}{2\mu}A^T A)x &\geq \lambda_{\min}(Z^T HZ)\|x_n\|_2^2 - 2\sigma_{\max}(H)\|x_p\|_2\|x_n\|_2 + \frac{1}{2\mu}\sigma_{\min}(A)\|x_p\|_2^2 \\
&\geq \lambda_{\min}(Z^T HZ)\|x_n\|_2^2 - 2\sigma_{\max}(H)\|x_p\|_2\sqrt{1 - \|x_p\|_2^2} + \frac{1}{2\mu}\sigma_{\min}(A)\|x_p\|_2^2 \\
&\geq \lambda_{\min}(Z^T HZ)(1 - \|x_p\|_2^2) - 2\sigma_{\max}(H)\|x_p\|_2 + \frac{1}{2\mu}\sigma_{\min}(A)\|x_p\|_2^2 \\
&\geq \left(\frac{1}{2\mu}\sigma_{\min}(A) - \lambda_{\min}(Z^T HZ)\right)\|x_p\|_2^2 - 2\sigma_{\max}(H)\|x_p\|_2 + \lambda_{\min}(Z^T HZ).
\end{aligned}$$

This is a quadratic in $\|x_p\|_2^2$ it is clearly positive everywhere if μ is chosen large enough such that

$$\left(\frac{1}{2\mu}\sigma_{\min}(A) - \lambda_{\min}(Z^T HZ)\right) > 0$$

and

$$\sigma_{\max}(H)^2 - \left(\frac{1}{2\mu}\sigma_{\min}(A) - \lambda_{\min}(Z^T HZ)\right)\lambda_{\min}(Z^T HZ) < 0.$$