Nonlinear Optimization I, Fall 2021
Extra Exercises (Not Graded)

Since we did not have homeworks covering conjugate gradient or quasi-newton, here are some exercises related to those that may help you getter understand the mechanisms behind them. This will not be graded and no solutions will be posted.

Exercise 1. For a symmetric positive definite matrix $A \in \mathbb{R}^{d\times d}$ and any collection of nonzero vectors $p_1, \dots p_n$, if these vectors are all A-conjugate, show they must be linearly independent. From this, conclude $n \leq d$ (that is, A has at most d conjugate directions).

Exercise 2. Implement the conjugate gradient method from lecture and run it on the randomized least squares problem considered in HW2. How does its performance compare to your three old implementations? Design two problem instances with very large and very small condition numbers for $A^T A$ and see how the convergence rate varies.

Exercise 3. Consider the quadratic minimization problem

$$\min_{x\in\mathbb{R}^d} g^T x + \frac{1}{2}x^T Bx \ .$$

(a) Propose (and prove) if and only if conditions on g and B for when a minimizer exist[1].

(b) Propose (and prove) if and only if conditions on g and B for when the minimizer is unique.

Exercise 4. Consider the trust region subproblem.

$$\min_{s} f + g^T s + \frac{1}{2}s^T Bs \qquad \text{subject to } \|s\|_2 \leq \Delta \tag{1}$$

with $g = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$ and $B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix}$.

(a) Using our exact characterization of the minimizers, compute a minimizing s^* when $\Delta = 2$.

(b) Using our exact characterization of the minimizers, compute a minimizing s^* when $\Delta = 5/12$.

(c) Letting $s^*(\Delta)$ denote the minimizer for some Δ, as $\Delta \to 0$, what happens to $s^*(\Delta)/\|s^*(\Delta)\|$?

[1]Hint: it is not sufficient for this objective to be convex. Why?