Assignment #7

Due Monday 11 November 2024, at the start of class

From the textbook[1] please read sections 2.5, 2.7, 11.2, 11.3, 11.4, 11.5, 12.1, 12.2, and Appendix B.5. (Sections 11.4 and 11.5 are important to know where we are going, but they are not needed for this Assignment.) Please also read the online slides *Steepest descent needs help* at

`bueler.github.io/opt/assets/slides/F24/sdneedshelp.pdf`

DO THE FOLLOWING EXERCISES from section 11.2, pages 361–364:

- Exercise 2.7
- Exercise 2.10
- Exercise 2.14

DO THE FOLLOWING EXERCISES from section 11.3, pages 369–371:

- Exercise 3.3 *Note: You are only asked to show that you have a* local *minimum.*
- Exercise 3.4
- Exercise 3.6 *Hint: Do* **P15** *first, and then use any result from it. You may assume Q is symmetric.*

Problem P15. *This problem collects together some basic facts about quadratic objective functions. These facts are assumed in Chapters 11 and 12. See Appendix B.5.*

Let $c \in \mathbb{R}^n$ be any vector, let $d \in \mathbb{R}$ be any number, and suppose $Q \in \mathbb{R}^{n\times n}$ is a symmetric positive definite matrix. Consider this quadratic function on $x \in \mathbb{R}^n$:

$$f(x) = \frac{1}{2}x^\top Qx - c^\top x + d.$$

a) Show that $\nabla f(x) = Qx - c$.

b) Compute the Hessian $\nabla^2 f(x)$.

c) Show that f is strictly convex. (*Hint. Easy. You may use facts from section 2.3.*)

d) Suppose p is a descent direction at x, so that $p^\top \nabla f(x) < 0$. Prove that the exact solution of the line search problem $\min_{\alpha>0} f(x+\alpha p)$ is

$$\alpha = \frac{-p^\top \nabla f(x)}{p^\top Qp}.$$

(*Hint. Define $g(\alpha) = f(x+\alpha p)$, expand it, and compute $g'(\alpha)$. Etc.*)

[1] Griva, Nash, and Sofer, *Linear and Nonlinear Optimization*, 2nd ed., SIAM Press 2009.

Problem P16. In the steepest descent slides I show a MATLAB implementation of steepest descent using back-tracking line search, something we will cover carefully in section 11.5. When the objective function $f(x)$ is quadratic then we can instead use the result in **P15 c)** to choose the step size.

a) Implement steepest descent with optimal step size for quadratic functions $f(x) = \frac{1}{2}x^\top Qx - c^\top x$:

```
function z = sdquad(x0,Q,c,tol)
```

Stop iterating when $\|\nabla f(x_k)\| < \texttt{tol}$. (*Hint. From the code in the slides, only small modifications are needed: Replace evaluations of f and ∇f by formulas for the quadratic case, and replace back-tracking by the result from* **P15 c)**.)

b) Use `sdquad()` to reproduce the result of Example 12.1 on pages 404–405 of the textbook. Specifically, you should get $k = 216$ iterations using $\texttt{tol} = 10^{-8}$.

c) Now change Q to

$$Q = \begin{pmatrix} 2.3 & 0.19 & -0.89 \\ 0.19 & 1.84 & 0.32 \\ -0.89 & 0.32 & 1.86 \end{pmatrix}$$

but keep the same c, x_0, and `tol` as in part **b)**. What is x_*? How many iterations does `sdquad()` need? Why is this problem easier than part **b)**? (*Hint. What does* `eig(Q)` *tell you?*)

Problem P17. *This problem extends Exercise 3.8 in section 11.3.*

Though section 2.5 fails to make this clear, a good definition of *superlinear convergence* is that, for sequences $\{x_k\}$ which converge to x_*,

$$\lim_{k\to\infty} \frac{\|e_{k+1}\|}{\|e_k\|} = 0$$

(Recall $e_k = x_k - x_*$.)

a) Let $\{x_k\}$ be a sequence which converges super-linearly to x_*. Show that

$$\lim_{k\to\infty} \frac{\|x_{k+1} - x_k\|}{\|x_k - x_*\|} = 1.$$

b) Explain why a super-linearly converging iterative algorithm for approximating some unknown number x_* can stop according to the criterion $\|x_{k+1} - x_k\| < \texttt{tol}$, for some user-supplied tolerance $\texttt{tol} > 0$. That is, explain why, with this criterion, the iterations only stops when x_k is close to x_*.

c) Should we apply the same stopping criterion from **b)** to iterations which converge only linearly?