Assignment #8

Due Wednesday 20 November 2024, at the start of class

From the textbook[1] please read sections 11.3, 11.4, 11.5, 12.1, 12.2, and 12.3. Significant goals of this Assignment are to understand the benefits of a line search (section 11.5), and the motivation for the quasi-Newton approach (section 12.3). Regarding several of the problems, you will need the exact line search formula for quadratic functions; it is the conclusion of Exercise 5.3 on page 386.

DO THE FOLLOWING EXERCISES from section 11.4, pages 374–375:

- Exercise 4.4

DO THE FOLLOWING EXERCISES from section 11.5, pages 385–391:

- Exercise 5.2 *Note: the sufficient decrease condition is at the bottom of page 377.*
- Exercise 5.5 *Hint: Define $F(\alpha) = f(x_k + \alpha p_k)$ and write down the condition at the minimizing α. This proof is short.*

DO THE FOLLOWING EXERCISES from section 12.2, pages 408–411:

- Exercise 2.3

DO THE FOLLOWING EXERCISES from section 12.3, pages 420–421:

- Exercise 3.1 *Hints: Write a code to do this; it is easier than doing even one step by hand (in my opinion). The symmetric rank-one update $B_{k+1} = B_k + \dots$ is on page 413, and again on page 414.*

Problem P18. Consider a one-variable problem

$$\min_{x\in\mathbb{R}} f(x)$$

where $f : \mathbb{R} \to \mathbb{R}$ is smooth. Recall that the Newton method for this problem solves $f'(x) = 0$ by the formulas $p_k = -f'(x_k)/f''(x_k)$ and $x_{k+1} = x_k + p_k$. The *secant method for minimization* only differs from the Newton method by replacing the second derivative with a difference quotient approximation based on the last two iterates:

$$f''(x_k) \approx \frac{f'(x_k) - f'(x_{k-1})}{x_k - x_{k-1}}.$$

Thus the secant method computes the step (search vector) by

$$p_k = -\frac{(x_k - x_{k-1})f'(x_k)}{f'(x_k) - f'(x_{k-1})},$$

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

and then it uses $x_{k+1} = x_k + p_k$ as before.

(a) Implement the secant method. Include a stopping criterion $|f'(x_k)| <$`tol`.

(b) Use your code to accurately solve $\min_{x\in\mathbb{R}} f(x)$, e.g. with `tol=` 10^{-8}, for the following functions and initial iterates:

i) $f(x) = x^3 - 2\sin x, \quad x_0 = 0, \quad x_1 = 1$
ii) $f(x) = 3x^4 - 4x^3 + 3x^2 - 6x, \quad x_0 = -1, \quad x_1 = 0$

(c) In part ii) above the exact minimum is at $x_* = 1$. Compute the errors $e_k = x_k - x_*$. Give evidence that the convergence is superlinear. Using the notation of section 2.5, what is your estimate of the rate (exponent) r?

Problem P19. Please read the discussion in Section 12.2 of why the steepest descent method is slow when applied to minimizing a quadratic function

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

The conclusion of Lemma 12.4 will make the most sense if you know that the level sets (= contours in $\mathbb{R}^2$) of $f(x)$ are generalized ellipses, and that the eccentricity of these sets is closely related to the condition number of Q. This question simply asks you to check this idea on a 2D example.

(a) Consider $f(x) = x_1^2 + 2x_2^2 - 3x_1$. What are Q and c in this case, and where is the minimizer x_*? Consider a contour $f(x) = \ell$ for some $\ell \in \mathbb{R}$. If nonempty, this contour is an ellipse; complete the square to put it in standard form

$$\frac{(x_1-\gamma)^2}{\alpha^2} + \frac{(x_2-\delta)^2}{\beta^2} = 1.$$

State $\alpha, \beta, \gamma, \delta$ in terms of ℓ. What is the ratio α/β?

(b) Use a contour plotter to plot some contours of the same function $f(x)$. Use `axis equal` or similar to make sure that the axes have the same scaling. What is the ratio of the largest to smallest dimensions of the ellipses you see? Compute $\mathrm{cond}(Q)$, and relate this value to the ellipses.