Problem 1 - Woodbury matrix identity

Let $A \in \mathbb{R}^{n \times n}$ and $C \in \mathbb{R}^{r \times r}$ be invertible matrices and let $U \in \mathbb{R}^{n \times r}$ and $V \in \mathbb{R}^{r \times n}$ be rectangular matrices such that $(C^{-1} + VA^{-1}U)$ is invertible. Show the following identity

$$(A + UCV)^{-1} = A^{-1} - A^{-1}U(C^{-1} + VA^{-1}U)^{-1}VA^{-1}.$$ 

Problem 2 - Modified Newton with constant stepsize

Consider running a Modified Newton’s Method which computes its search direction at each iteration as

$$p_k = \text{argmin} \left\{ g_k^T p + \frac{1}{2} p^T B_k p \right\} = -B_k^{-1} g_k$$

where $g_k = \nabla f(x_k)$ and $B_k$ is some positive definite matrix selected at each iteration. Assume that every $B_k$ has eigenvalues at least $\lambda_{\text{min}}>0$ and at most $\lambda_{\text{max}}<\infty$. In this question, you will analyze this method’s convergence when using a constant stepsize $x_{k+1} = x_k + \alpha p_k$.

(a) For any function $f$ with $L$-Lipschitz gradient, show that for all $x_k$, the following relation holds

$$f(x_{k+1}) \leq f(x_k) + \alpha g_k^T p_k + \frac{L\alpha^2}{2} \|p_k\|^2$$

and deduce that

$$f(x_{k+1}) \leq f(x_k) - \left( \frac{\alpha}{\lambda_{\text{max}}} - \frac{L\alpha^2}{2\lambda_{\text{min}}} \right) \|g_k\|^2.$$ 

(b) Consider the quadratic $\left( \frac{\alpha}{\lambda_{\text{max}}} - \frac{L\alpha^2}{2\lambda_{\text{min}}} \right) \|g_k\|^2$ in terms of $\alpha$ above. What value of $\alpha$ maximizes this amount? Is this maximum quantity positive or negative? Is the maximizing value of $\alpha$ positive or negative?

(c) Using this maximizing value of $\alpha$ as a constant stepsize choice, derive a convergence guarantee for this method of the form

$$\min_{i \leq k} \|\nabla f(x_i)\| \leq \frac{M}{\sqrt{k+1}}$$

for some constant $M$ depending on $\lambda_{\text{min}}$, $\lambda_{\text{max}}$, $L$, and $f(x_0) - \min f$.

(d) What can you say about your algorithm and guarantee under the choice of $B_k = LI$?
Problem 3 - BFGS gives descent

Consider running a Quasi-Newton Method where our stepsize is selected to ensure $y_k^T s_k > 0$.

(a) Supposing $B_k$ is positive definite, show the BFGS update produces a positive definite $B_{k+1}$.

(b) Now that you have guaranteed the update from BFGS is invertible (by showing it is positive definite above), calculate its inverse using the Woodbury matrix identity from Problem 1.

Problem 4 - A scary function for GD

Consider the two-dimensional Rosenbrock function (a common example used to show the steepest descent method slowly converges):

$$\min_{x \in \mathbb{R}^2} (1 - x_1)^2 + 100(x_2 - x_1^2)^2.$$  

Note this problem globally minimizes at $x^* = (1, 1)$.

(a) Implement and run gradient descent on this problem initialized at $x_0 = (0, 0)$ for 100 iterations using an exact linesearch$^1$. Print out the distance to the solution $\|x_k - x^*\|$ at each iteration.

(b) Implement and run Newton’s Method on this problem for 100 iterations initialized at $(0, 0)$. Print out $\|x_k - x^*\|$ at each iteration.

(c) Implement and run the BFGS Quasi-Newton Method for 100 iterations initialized at $(0, 0)$ using an exact linesearch. Print out $\|x_k - x^*\|$ at each iteration.

---

$^1$That is, given a direction to search $p_k$, select $\alpha_k$ minimizing $f(x_k - \alpha_k p_k)$. Here this corresponds to minimizing a single variable, degree four polynomial.