ACM41020 Assignment 3

Optimisation

DUE: 10AM FRIDAY 8TH NOVEMBER 2024

At several points in this assignment we will consider the particular case where

$$S = \begin{bmatrix} 5 & -3 \\ -3 & 5 \end{bmatrix}, \qquad A = \begin{bmatrix} 1 \\ 3 \end{bmatrix}.$$

1. **Constrained Optimisation**
   We wish to minimise the quadratic function $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T S \mathbf{x}$ subject to $A^T\mathbf{x} = b$.
   a) What is the Lagrangian $L(\mathbf{x}, \lambda)$ for this problem?
   b) What are the equations "derivative of L = zero"?
   c) Solve those equations to find the location $\mathbf{x}^* = (x_1^*, x_2^*)$ of the minimum $F_{\min} = F(\mathbf{x}^*)$. Solve also for the multiplier λ^*.
   d) Draw or plot the constraint line $A^T\mathbf{x} = b$ tangent to the ellipse $F(\mathbf{x}) = F_{\min}$.
   e) Verify that the derivative of the minimum cost is $\dfrac{\partial F_{\min}}{\partial b} = \lambda^*$.

2. **Unconstrained Optimisation Using Gradient Descent**
   a) Derive an expression in terms of S and $\mathbf{x}$ for the optimal step size for gradient descent applied to find the unconstrained minimum of a quadratic function $F(\mathbf{x}) = \frac{1}{2}\mathbf{x}^T S \mathbf{x}$ with a generic symmetric and positive-definite matrix S.
   b) By considering the worst case scenario (worst $\mathbf{x}_i$) for the decrease in the error $e_k = ||\mathbf{x}_k - \mathbf{x}^*||$ from one step to the next, derive an expression for the optimal step size in that case in terms of the eigenvalues of S.
   c) What condition must $\mathbf{x}_i$ satisfy to be in this worst-case scenario?
   d) Working with the specific matrix S above and starting from $\mathbf{x}_0 = (3,5)$ compute six steps using gradient descent with the optimal step size.
   e) Sketch/plot the contours defined by the ellipses $F(\mathbf{x}) = F(\mathbf{x}_i)$ for $i = 0, 1, \ldots, 6$. On the same plot indicate the points $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$, $\mathbf{x}_5$, $\mathbf{x}_6$ obtained in part (d) and the lines defined by the set of worst-case points identified in part (c).
   f) Compute the error e_k in each step. By considering how the error decreases over 3 steps verify that the iteration is converging linearly to the solution $\mathbf{x}^*$.

3. **Gradient Descent with Momentum**
   a) Working with the matrix S above and starting from $\mathbf{x}_0 = (3,5)$ compute six steps using gradient descent with momentum. Use the values for the optimal step size α and momentum parameter β in terms of the eigenvalues of S as given in the lectures.
   b) Sketch/plot the contours defined by the ellipses $F(\mathbf{x}) = F(\mathbf{x}_i)$ for $i = 0, 1, \ldots, 6$. On the same plot indicate the points $\mathbf{x}_0$, $\mathbf{x}_1$, $\mathbf{x}_2$, $\mathbf{x}_3$, $\mathbf{x}_4$, $\mathbf{x}_5$, $\mathbf{x}_6$.
   c) Compute the error e_k in each step.
   d) Briefly (1-2 sentence) compare and contrast the results with and without momentum.

Note: The questions in this assignment can be completed with by-hand calculations, but if you wish you may use a computer to help with any number-crunching.