NUMERICAL OPTIMIZATION

Sheet 8: Conjugate gradient method

Exercise one Execute the conjugate gradient method to solve the system $Ax = b$ for

$$A = \begin{pmatrix} 7 & -3 & 1 & -1 \\ -3 & 7 & -1 & 1 \\ 1 & -1 & 7 & -3 \\ -1 & 1 & -3 & 7 \end{pmatrix}$$

and $b = (4, 0, 8, 4)$, starting from the initial point $x_0 = (0, 0, 0, 0)$.

List all the residuals r_i and directions p_i that you see along the way.

Note: If you wish to review the algorithm, it is listed as Algorithm 5.2 (CG) in Nocedal-Wright. I have checked this exercise and all the numbers on the way including the optimal solution are quite well-behaved.

Exercise two Leftovers from the lecture:

1. In the proof of the first theorem (Theorem 5.1 from Nocedal-Wright) we needed the following equality: $A(x^* - x_k) = b - Ax_k = -r_k$. Prove that this equality indeed holds.
2. Prove the useful recurrence for the upcoming residual direction as a combination of the previous residual direction and Ap_k:

$$r_{k+1} = r_k + \alpha_k A p_k.$$

Exercise three Review the proof of Theorem 5.3 from Nocedal-Wright that we have investigated at the lecture. In particular, we have not proven point (1), which is called (5.16) in Nocedal-Wright:

$$\forall i \in \{0, 1, \dots, k-1\} \colon r_k^T r_i = 0. \qquad (5.16)$$

Explain the proof of this fact and check that your residuals from exercise one satisfy it.

Exercise four

Part 1. Recall the *Spectral theorem* from linear algebra and use it to prove the following:

Let A be a positive definite matrix and x any vector. Prove that $x = \sum_{i=1}^{l} y_i$ where each y_i is an eigenvector of A corresponding to eigenvalue λ_i, and l is a number of *distinct* eigenvalues.

Note: Observe that y_i have no coefficients in front of themselves, and that we are not talking about a basis – each x can have its own set y_i. Think about what the statement says about the identity matrix, which is positive definite.

Part 2. Suppose we are back to our conjugate gradient method solving $Ax = b$ for A positive definite, which we already know is equivalent to $\min \frac{1}{2} x^T A x - b^T x$, starting from the point x_0.

Using Part 1, we can write $x^* - x_0 = \sum_{i=1}^{l} y_i$. Let $W = \operatorname{span}\{y_1, y_2, \dots, y_l\}$ be the linear subspace generated by y_i. Prove, by induction over k, that the conjugate gradient method never leaves the affine subspace $x_0 + W$.

Note: As with many other exercises of Numerical optimization, this sounds tougher than it really is. We already know the expressions for computing r_{k+1}, p_{k+1} and x_{k+1}, so we just need to check that we start and stay in the subspace $x_0 + W$.

Exercise five Solve the problem $Ax = b$ for

$$A = \begin{pmatrix} 1 & 0 & 1 & -2 \\ 0 & 1 & -2 & 0 \\ 1 & 2 & 1 & 0 \\ 2 & 1 & 3 & -2 \end{pmatrix}$$

and $b = (6, 3, -2, 0)$ using Gaussian elimination.

Also check for A whether A and all its leading principal minors have positive determinant.

Exercise six Try to solve the problem $Ax = b$ using the conjugate gradient method. How many iterations did it take for you and why?

Note: For this exercise, you are allowed and advised to implement the conjugate gradient method and run the code to get the answer. Again, I have checked and it is quite easy using Numpy and `numpy.dot` for scalar product and matrix/vector multiplication.