Question 14

The implemented code for this exercise is based on the lecture slide about the conjugate gradient method. The code also contains comments explaining what is done in each steps.

The code take the matrix A(symmetric positive definite) , vector b and X0 as the initial starting point, returning X as the solution of the system of linear equations. Based on the theorem discussed in the lecture, the conjugate gradient method is a conjugate directions method. As a result, the algorithm is terminated at most n steps resulting the exact minimizer of f (solution of Ax=b).

The code also checks whether the gradient is equal to zero in each iteration. As a consequence, the code is guaranteed to terminate within n steps(less steps if the gradient reaches zero in earlier steps).

Definition (conjugate gradient method)

Let

1. $A \in \mathbb{R}^{n\times n}$ be a symmetric positive definite matrix, $b \in \mathbb{R}^n$, $f : \mathbb{R}^n \to \mathbb{R}$ the associated quadratic form,
2. $x_0 \in \mathbb{R}^n$ an arbitrary starting point.

Initialization: $d_0 = -g_0 = -\nabla f(x_0) = b - Ax_0$.
Iteration ($k = 0, \ldots, n-1$, while $g_k \neq 0$):

$$x_{k+1} = x_k + \gamma_k d_k, \quad \gamma_k = -\frac{d_k^T g_k}{d_k^T A d_k},$$

$$d_{k+1} = -g_{k+1} + \beta_k d_k, \quad \beta_k = \frac{d_k^T A g_{k+1}}{d_k^T A d_k},$$

where g_k is the gradient of f at x_k: $g_k = \nabla f(x_k) = Ax_k - b$.

(The iteration terminates at step k, if $g_k = \nabla f(x_k) = 0$).

Note :

Out objective function is a quadratic function as follows:

f(x)=$\frac{1}{2} x^T$Ax-b^Tx

and the derivative of (f) is:

df(x)=Ax-b

which is the solution of the system of linear equations Ax=b.

so the minimum of f is the solution of the system of linear equations and the minimum is found after running all the iterations.

Note: the theorem mentioned in the lecture slides also guarantees the convergence of the method in finite steps.

Theorem

The *conjugate gradient method* is a conjugate directions method. Assume that it does not terminate at step k, then

1. The gradients $g_0, \dots, g_k$ at points $x_0, \dots, x_k$ are nonzero and
$$\mathrm{Lin}\{g_0, g_1, \dots, g_k\} = \mathrm{Lin}\{g_0, Ag_0, \dots, A^k g_0\}.$$
2. The direction $d_0, \dots, d_k$ are nonzero and
$$\mathrm{Lin}\{d_0, d_1, \dots, d_k\} = \mathrm{Lin}\{g_0, Ag_0, \dots, A^k g_0\}.$$
3. The direction $d_0, \dots, d_k$ are A-orthogonal.
4. The algorithm is terminated at most n steps resulting the exact minimizer of f (solution of Ax=b).