Lesson 18

1. What is the major difference between the method of Conjugate Directions and the method of Conjugate Gradients?

   The method of conjugate gradients is the method of conjugate directions with the modification that the search directions are set up by conjugation of the residuals. Since the residuals have the property that they are orthogonal to the previous search directions, construction of the search directions from the residuals is guaranteed to ensure a new linearly independent search direction. The only exception is when the residual becomes zero. This is not a problem because it means that at that point the residual has been found.

2. How is the process of finding a new search direction for the method of conjugate gradients simplified?

   In the conjugate gradient method, finding the new search direction $\mathbf{d}_{i+1}$ from $\mathbf{r}_{i+1}$ simple since $\mathbf{r}_{i+1}$ is orthogonal to $\mathbf{A}D_i$ and hence $\mathbf{A}$-orthogonal to $D_i = \{\mathbf{d}_0, \mathbf{d}_1, \ldots \mathbf{d}_{i-1}\}$ i.e. to all previous directions except $\mathbf{d}$i. Thus the Gramm –Schmidt procedure need only ensure $\mathbf{A}$-orthogonality with $\mathbf{d}_i$

   Because of this it is no longer necessary to store the old search vectors to ensure conjugacy with the old search directions. This not only reduces storage but also drastically curtails the computations necessary to find the new search directions.

3. Will the conjugate gradient method always converge to the minimum for a *n*-dimensional quadratic form in *n* iterations?

   By construction, for a *n*-dimensional problem, the conjugate gradient algorithm is bound to converge to the minimum of a quadratic form in *n* iterations. However floating point errors accumulate with number of iterations, causing the residuals to lose orthogonality and hence the search directions to loose **A**-orthogonality. Hence faster convergence by reducing the accumulation of errors improves the accuracy and optimality of the CG method.

4. What is the criterion for the conjugate gradient method to have good convergence properties?

   It has been seen that the convergence of the conjugate gradient method is faster when the eigenvalues of **A** are clustered together. This is likely to happen when the condition number of **A** is small.

5. What is meant by pre-conditioning? How does pre-conditioning improving the convergence of the conjugate gradient method?

It is clear that if we can improve the condition number of matrix **A**, we can improve the rate of convergence. This can be done for instance by a process known as preconditioning which involves scaling **A** with the inverse of symmetric positive definite matrix **M**. If the condition number of $\mathbf{M}^{-1}\mathbf{A}$ is smaller than the condition number of **A**, then we can iteratively solve more quickly than the original problem.