# Lesson 17

1. Why is the Conjugate Directions method an improvement over the Steepest Descent method?

   In the steepest descent method we often need to take steps in the same direction as earlier steps. It would be more efficient if we do not have to retrace our steps. This can be done if we make sure that the search directions are orthogonal to each other – and in each search direction we take steps of exactly the right length. Then in an n-dimensional space, after n such orthogonal steps of exactly right length – we are sure to reach the minimum. The idea is that during a particular step in the iteration, we exhaust the error in that direction and we ensure that the following step introduces no additional errors in the direction of the previous step. The conjugate directions method tries to achieve this and is thus an improvement over the Steepest Descent method.

2. Are the search directions in the Conjugate Directions method truly orthogonal to each other?

   To be of practicable, the conjugate search directions cannot be made truly orthogonal to each other. Instead, they are **A**-orthogonal to each other i.e. they are orthogonal to each other with respect to the coefficient matrix **A**: $\mathbf{d}_i^T \mathbf{A} \mathbf{d}_{i-1} = 0$ where $\mathbf{d}_i$ and $\mathbf{d}_{i-1}$ are successive conjugate directions.

3. How many steps are required in the Conjugate Directions method to minimize a *n*-dimensional quadratic form?

   The **A** orthogonality of the conjugate search directions ensures that a *n*-dimensional quadratic form is minimized after n steps.

4. How are the search directions in the Conjugate Directions method set up?

   The Gramm-Schmidt procedure, which basically sets up *n* orthogonal basis vectors from a set of *n* linearly independent vectors, is used to set up the conjugate directions.

5. What is the "best solution property" of the method of conjugate directions?

   An important property of the method of conjugate directions is that it finds the best solution within the bounds of the space where it has been allowed to operate. This follows from the fact that at each step the conjugate direction is so chosen that it minimizes the $L_2$ norm of the error with respect to the coefficient matrix **A** for that step.

6. What does orthogonality of the residuals mean in the context of the method of conjugate directions?

In the method of conjugate directions the residual $\mathbf{r}_i$ at step $i$ is orthogonal (not **A**-orthogonal) to the old search directions ie. $\mathbf{r}_\mathrm{i}$ is orthogonal to the space spanned by $\{\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \ldots..\mathbf{d}_{i-1}\}$ where $\{\mathbf{d}_0, \mathbf{d}_1, \mathbf{d}_2, \ldots..\mathbf{d}_{i-1}\}$ are the search directions upto the $i^{\mathrm{th}}$ step. This property is used in the formulation of the method of Conjugate Gradients.