3.3. SECOND DERIVATIVE METHODS

3.3.1. NEWTON'S METHOD

Like Newton's method in univariate quadratic interpolation, f is approximated at $\mathbf{x_k}$ with Taylor's polynomial of second order:

$$f(\boldsymbol{x}_k+\boldsymbol{p}) \approx f(\boldsymbol{x}_k) + \boldsymbol{p}^T\nabla f(\boldsymbol{x}_k) + \frac{1}{2}\boldsymbol{p}^T H(\boldsymbol{x}_k)\boldsymbol{p}$$

The minimum of the right hand side is achieved by minimizing the function

$$\Phi(\boldsymbol{p}) = \boldsymbol{p}^T\nabla f(\boldsymbol{x}_k) + \frac{1}{2}\boldsymbol{p}^T H(\boldsymbol{x}_k)\boldsymbol{p}$$

$$\Rightarrow\ \nabla\Phi(\boldsymbol{p}) = \nabla f(\boldsymbol{x}_k) + H(\boldsymbol{x}_k)\boldsymbol{p} = \mathbf{0}.$$

The displacement vector $\mathbf{p_k}$ is solved from the linear equations

$H(\mathbf{x_k})\mathbf{p_k} = -\nabla f(\mathbf{x_k}) \quad \Leftrightarrow \quad \mathbf{p_k} = -H(\mathbf{x_k})^{-1}\nabla f(\mathbf{x_k})$ (*the Newton direction*)

Newton's iteration formula:

$$\boxed{\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - H(\boldsymbol{x}_k)^{-1}\nabla f(\boldsymbol{x}_k)}$$

If f is a quadratic function with a positive definite Hessian, the model is accurate and the method reaches the minimum with one iteration, from any starting point.

In general, a line search (a step length procedure) must be included because a step of unity along the Newton direction will not necessarily reduce f even if it is a step to the minimum of the approximating quadratic function. In *the modified Newton's method* the iteration formula is

$$\boxed{\boldsymbol{x}_{k+1} = \boldsymbol{x}_k - \alpha_k H(x_k)^{-1}\nabla f(\boldsymbol{x}_k)}$$

When the Hessian matrices $H(\mathbf{x_k})$ are positive definite and normal step length conditions are valid, the method is a descent method and converges quadratically to the minimum point $\mathbf{x^*}$.

A constant step length 1 can be used, but then the convergence or failure of the method is dependent on the accuracy of the quadratic model and and also on the distance from the starting point to the minimum point.

Although the modified Newton's method is reliable and efficient when first and second derivatives are easily calculated, its disadvantage is the need to solve the system of equations involving the Hessian at each iteration.

Special strategies are needed for the case of an indefinite Hessian. For example, if $\mathbf{x_k}$ is a *saddle point*, i.e. a stationary point where $H(\mathbf{x_k})$ is indefinite, the direction $\mathbf{p_k}$ can be chosen as a direction of negative curvature: such that

$$\mathbf{p_k}^T H(\mathbf{x_k})\mathbf{p_k} < 0.$$

When $H(\mathbf{x_k})$ is indefinite, such a direction must exist and f can be decreased by taking a step in that direction.

When $H(\mathbf{x_k})$ is singular and $\nabla f(\mathbf{x_k}) \neq \mathbf{0}$, a steepest descent direction can be used.

<u>Example 3.3.</u> Minimize $f(\mathbf{x}) = (x_1 - 2)^4 + (x_1 - 2x_2)^2$ with the Newton's method starting from the point $\mathbf{x_0} = (0, 3)$. The exact minimum is $f^* = f(2, 1) = 0$.