Gradient solvers : February 27, 2021

- Steepest decent : Looking only at the gradient vector **g**. We go in it's direction such that $x_{i+1} = x_i - S_i\hat{g}_i$. So we move in the direction where S is the step length to the point where W is minimum in that direction
- Relaxed steepest decent : We have a relaxation in the step length with hS, where h is a coefficient on the step length
- Runge-Kutta : The decent is only true at the inital point , and can diverge. Therefore we use different averages of the gradient.
- Newton Rhapson method: The gradient is linearized using Taylor series such that $g_{k+1} = g_k + \frac{d^2 w}{dx^2}|_x \Delta x$. Here we need to find the derivative of the gradient (a vector) giving us the hessian matrix. Sometimes this can be semi definite , so we may need to decrease the step size.
- Fletcher-Powell : We avoid computing the inverse of the hessian, replacing it by a sequence of positive definite matrices.
- Conjugate gradients : Originally developed for a solution of a system of lienar equations having positive definite matrix coefficnet. Or $\hat{g}_i{}^T H\hat{g}_i = 0$. FIX ME. Still have to read about this