Trust Region Method

As opposed to line search methods in which we generate a search direction and choose a suitable step length, trust-region methods work by generating a region around the current iterate in which we minimize the model function. To obtain the model function we take the quadratic approximation of the function about the given point i.e.

$$f(x_k + p) = f_k + g_k^T p + \frac{1}{2} p^T \nabla^2 f(x_k + tp) p$$

where $t \in (0, 1)$, step $p = (x - x_k)$, g_k is the gradient of f at x_k. If we approximate the hessian with a symmetric matrix B_k we can write the model function about x_k as

$$m_k(p) = f_k + g_k^T p + \frac{1}{2} p^T B_k p$$

1 Trust Region Approach

Let $\Delta_k > 0$ be the trust radius and Ω_k be the region of trust where

$$\Omega_k = \{x : \|x - x_k\| \leq \Delta_k\}$$

Our aim is to minimize the following constrained optimization problem

$$\min_p m_k(p) \text{ s.t. } \|p_k\| \leq \Delta_k$$

where the constraint $\|p_k\| \leq \Delta_k$ essentially means $x \in \Omega_k$.

Algorithm 1 To Determine Δ_{k+1} and R_k

Require: Δ_k, x_k, p_k
  $R_k = \frac{f(x_k) - f(x_k + p_k)}{m_k(0) - m_k(p_k)}$
  if $R_k < 0.25$ **then**
    $\Delta_{k+1} = \Delta_k / 4$
  else if $R_k > 0.75$ and $\|p_k\| == \Delta_k$ **then**
    $\Delta_{k+1} = 2 * \Delta_k$
  else
    $\Delta_{k+1} = \Delta_k$
  end if
  return Δ^{k+1}, R^k

For the solution of the above optimization problem we need to determine Δ_{k+1}. Choice of the trust radius is one of the crucial steps in trust-region methods. If chosen too small we may miss out an opportunity to move closer to the minima, whereas a too large step will take us farther away from the minima. The choice is based on the agreement between the decrease in model function m_k and the objective function f in the previous iteration. Algorithm 1 describes the process of obtaining Δ_{k+1}

For algorithm 1 we require p_k and hence we first need to solve the subproblem. It's characterized by the following theorem.

Theorem 1. *p^* is a global solution of the trust region problem*

$$\min_p m_k(p) = f + g^T p + \frac{1}{2} p^T B p \ \ s.t. \ \ \|p_k\| \leq \Delta_k$$

iff p^* is feasible and there exists a scalar $\lambda \geq 0$ such that the following conditions are satisfied

$$\begin{aligned}&(B + \lambda I)p^* = -g\\&\lambda(\Delta - \|p^*\|) = 0\\&(B + \lambda I) \textit{is positive semi-definite}\end{aligned}$$

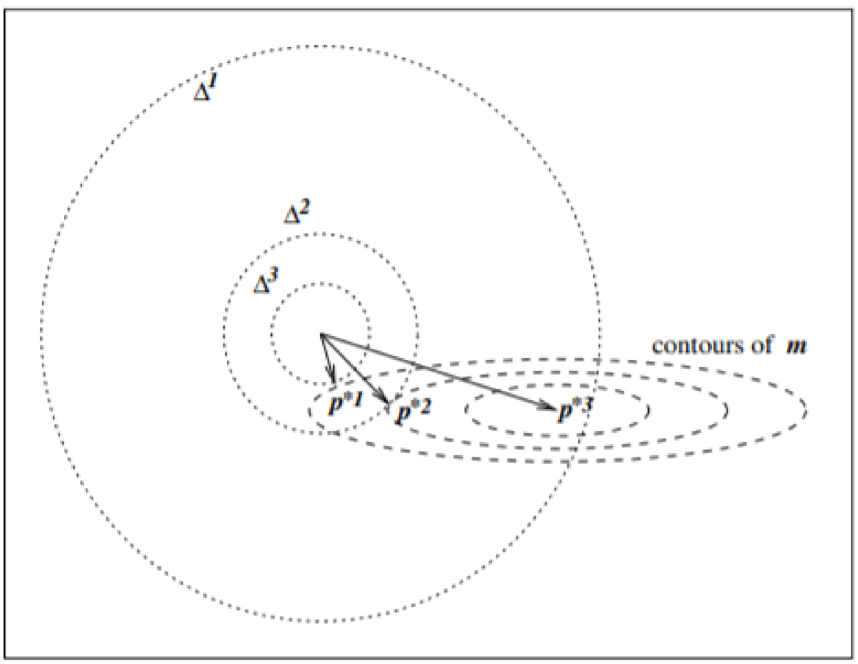

Figure 1: Solution of trust-region sub-problem for different radii (src: Nocedal & Wright)

In the above theorem the second condition is the complementarity condition where either $\lambda = 0$ or $(\Delta - \|p^*\|) = 0$. When p^* is within the trust region $(\Delta - \|p^*\|) > 0$ and therefore λ should be 0. Hence, condition 1 becomes $Bp^* = -g$ same as the unconstrained minimizer of m_k. In other cases the optimal value of p is at the boundary of the trust region and therefore $(\Delta - \|p^*\|) = 0$ so λ can be greater that 0.

2 Method to solve the subproblem: Dogleg Method

Let B_k be the Hessian or an approximation to the Hessian and if B_k is positive definite then the Dogleg method approximates the optimal trajectory $p(\Delta)$ by a path connected by two line segments one along the unconstrained minimizer along the steepest direction and the other along p^U to p^B, where $p^U = -\frac{g^T g}{g^T B g} g$ is the steepest direction and p^B is the optimal solution of the quadratic model. The approximate solution to the optimal trajectory is given by the following equations

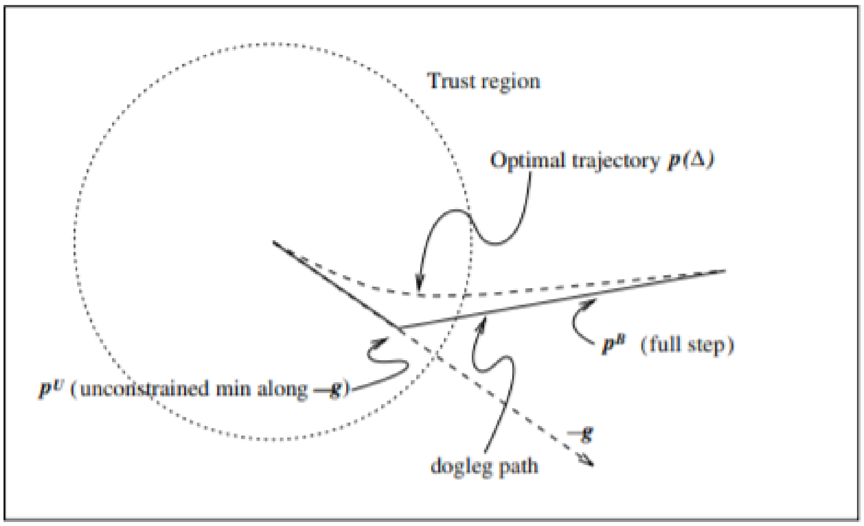

Figure 2: Exact trajectory and Dogleg approximation (src: Nocedal & Wright)

$$
\begin{aligned}
p(\tau) &= \tau p^U, && \text{if } 0 \leq \tau \leq 1 \\
p(\tau) &= p^U + (\tau - 1)(p^B - p^U), && \text{if } 1 \leq \tau \leq 2
\end{aligned}
$$