

Numerical Optimisation:
Trust Region Methods

Assignment Project Exam Help

Marta M. Betcke

`m.betcke@ucl.ac.uk`,

Kiko Rullan

`f.rullan@cs.ucl.ac.uk`

<https://powcoder.com>

Add WeChat powcoder

Department of Computer Science,
Centre for Medical Imaging Computing,
Centre for Inverse Problems,
University College London

Assignment 3

Recall the constraint minimisation problem for the trust region method with quadratic approximation of the function:

$$\min m(p) = f(x_k) + g^T p + \frac{1}{2} p^T B p \quad \text{s.t. } \|p\| \leq \Delta$$

Let us consider S the subspace spanned by g and $B^{-1}g$.

$$S = \text{span}(g, B^{-1}g)$$

We can take an orthonormal basis V in S and express p as a linear combination of this basis via:

$$p = Va$$

Now consider the minimisation problem in terms of this basis:

$$\min m_v(a) = f(x_k) + g_v^T a + \frac{1}{2} a^T B_v a \quad \text{s.t. } \|a\| \leq \Delta,$$

Assignment Project Exam Help

where we have used the fact that $V^T V = I$. As long as V has a full rank (g and $B^{-1}g$ are not collinear), if B is s.p.d. so is B_v .

Note that if $g = cB^{-1}g$ the problem becomes 1D.

<https://powcoder.com>

To solve the projected model m_v subject to $\|p\| \leq \Delta$ we make use of Theorem 4.1 Nocedal Wiright. From this theorem for m_v we have that a minimizes m_v s.t. $\|a\| \leq \Delta$ iff

Add WeChat powcoder

$$\begin{cases} (B_v + \lambda I) a = -g_v, & \lambda \geq 0 \\ \lambda(\Delta - \|a\|) = 0, \\ (B_v + \lambda I) \text{ is s.p.d.} \end{cases} \quad (1)$$

This gives two cases:

- $\lambda = 0$ and $\|a\| < \Delta$. The unconstrained solution is inside the trust region. Then the first equation becomes:

$$B_V a = -g_V \Rightarrow a = -B_V^{-1} g_V$$

- $\lambda \geq 0$ and $\|a\| = \Delta$. The constraint is active. Then we can solve the first equation:

$$a = -(B_V + \lambda I)^{-1} g_V$$

The additional equation is provided by the constraint:

$$\|a\| = \Delta$$

To solve this system we make use of eigendecomposition of B_V :

$$B_V = Q^T D Q \quad \text{with } Q \text{ orthonormal}$$

Then we have:

$$Qa = -(D + \lambda I)^{-1} Qg_v$$

and realise that $(Qa)^T(Qa) = a^T Q^T Q a = a^T a$. We denote

$Q_{a,i} = (Qa)_i$ and $Q_{g,i} = (Qg_v)_i$. For i -th element on Qa :

$$Q_{a,i} = -\frac{1}{(d_i + \lambda)} Q_{g,i},$$

with d_i the i -th element in the diagonal of D . Now, substituting Q_a into $\|Q_a\|^2 = Q_{a,1}^2 + Q_{a,2}^2 = \Delta^2$ we obtain:

$$\frac{Q_{g,1}^2}{(d_1 + \lambda)^2} + \frac{Q_{g,2}^2}{(d_2 + \lambda)^2} = \Delta^2$$

which we can transform to a 4th degree polynomial in λ assuming that $d_i + \lambda > 0$.