LM Introduction

spc

LevenbergMarquardt method is a kind of trust-region method in Numerical Optimization. It can be used to solve nonlinear least square problem. In effect, trust-region menthod choose the direction and length of the step simultaneously. For a nonlinear least square problem:

$$F(x) = \frac{1}{2}\|\boldsymbol{f}(x)\|^2 \tag{1}$$

Based on the Taylor-series expansion of $\boldsymbol{f}(x)$, we can get

$$\boldsymbol{f}(x_k + p) = \boldsymbol{f}(x_k) + \boldsymbol{J}(x_k)p + O(p^T p) \tag{2}$$

$$F(x_k + p) \approx \frac{1}{2}\|\boldsymbol{f}(x_k)\|^2 + p^T \boldsymbol{J}^T(x_k)\boldsymbol{f}(x_k) + \frac{1}{2}p^T \boldsymbol{f}^T(x_k)\boldsymbol{f}(x_k)p = m_k(p) \tag{3}$$

To obtain each step, we seek a solution of the subproblem

$$\min_{p \in R^n} m_k(p) = \frac{1}{2}\|\boldsymbol{f}(x_k) + \boldsymbol{J}(x_k)p\|^2 \qquad s.t. \|D_k p\|^2 \leq \mu_k \tag{4}$$

where $\mu_k > 0$ is the trust-region radius and we limit step p to an ellipse. Using the Lagrange Multiplier method, we can turn it into an unconstrained Optimization problem

$$\min_{p \in R^n} = \frac{1}{2}\|\boldsymbol{f}(x_k) + \boldsymbol{J}(x_k)p\|^2 + \frac{\lambda}{2}\|D_k p\|^2 \tag{5}$$

In discussing this matter, we sometimes drop the iteration subscript k of (5). By taking the derivative of it with respect to p, we can obtian the solution p^* of (5) satisfies

$$\left(\boldsymbol{J}^T\boldsymbol{J} + \lambda D^T D\right) p^* = -\boldsymbol{J}^T\boldsymbol{f} \tag{6}$$

Let $H = \boldsymbol{J}^T\boldsymbol{J}$ and $g = -\boldsymbol{J}^T\boldsymbol{f}$, (6) can also be written as

$$\left(H + \lambda D^T D\right) p^* = g \tag{7}$$

On the one hand, when λ is small, the approximation effect of the quadratic model is good and the LM method is closer to Gauss-Newton menthod. On the other hand, when λ is large, the LM method is closer to first order gradient descent method.

Strategy for choosing the trust-region radius μ_k at each iteration is one of the key problems in a trust-region algorithm. We base this choice on on the agreement between the objctive function f and model function m_k at previous iterations

$$\rho_k = \frac{F(x_k) - F(x_k + p_k)}{m_k(0) - m_k(p_k)} \tag{8}$$

Note that since the step p_k is obtained by minimizing the model m_k over a region that includes $p = 0$, the denominator will always be nonnegative. Hence, if ρ_k is negative, the new objective value is greater than the current value, so the step must be rejected. On the other hand, if ρ_k is close to 1, the agreement between m_k and f is good, so it is safe expand the trust region for the next iteration. If ρ_k is close to zero or negative, we shrink the trust region at next iteration, but if it is positive but significantly smaller than 1, we do not alter the trust region. The whole descriptions of the LM method is as follows:

Algorithm 1 Levenberg Marquardt method

Input: x_k, ρ_l, ρ_h, λ_k, μ_k, lc, M, ϵ_1, ϵ_2.
Initialize: $k = 0$, $\rho_l = 0.25$, $\rho_h = 0.75$, $\lambda_0 = 1$, $\mu_0 = 2$, $lc = 0.75$.

while $k < M$ **do**
    Compute J_k, H_k, g_k and obtain p_k.
    Evaluate ρ_k. $\rho_k = \frac{F(x_k) - F(x_k + p_k)}{m_k(0) - m_k(p_k)}$
    if $\rho_k > \rho_h$ **then**
        $\lambda_{k+1} = \frac{1}{2}\lambda_k$;
        if $\lambda_{k+1} < lc$ **then**
            $\lambda_{k+1} = 0$;
        end if
    else if $\rho_k < \rho_l$ **then**
        Find a new μ_k.
        $\lambda_{k+1} = \lambda_k * \mu_k$;
    end if
    if $F(x_k) > F(x_k + p_k)$ **then**
        $x_{k+1} = x_k + p_k$;
        $\mu_{k+1} = \mu_k$;
    end if
    if $\|p_k\|_\infty < \epsilon_1$ and $\|\boldsymbol{f}_k\|_\infty < \epsilon_2$ **then**
        break;
    end if
end while
