Levenberg-Marquardt Minimization

Jörn Wilms
Remeis-Sternwarte & ECAP
Universität Erlangen-Nürnberg

http://pulsar.sternwarte.uni-erlangen.de/wilms/
joern.wilms@sternwarte.uni-erlangen.de

FRIEDRICH-ALEXANDER
UNIVERSITÄT
ERLANGEN-NÜRNBERG

NATURWISSENSCHAFTLICHE
FAKULTÄT

ERLANGEN CENTRE
FOR ASTROPARTICLE
PHYSICS

Introduction

The fitting problem:

- Given the linear model

$$N_{\text{ph, predicted}}(c;\rho) = \Delta T \cdot \sum_{i=0}^{n_{\text{ch}}} A(E_i) \cdot R(c,i) \cdot M(E_i;\rho) \cdot \Delta E_i + N_{\text{background}}(c)$$

$$\forall c \in \{1, 2, \ldots, n_{\text{en}}\} \quad (1)$$

  where $M(E_i;\rho)$ is the spectral model, which depends on parameters $\mathbf{x}$, and

- given a fit statistics

$$S^2(\mathbf{x}) = f\left(N_{\text{ph, measured}}, N_{\text{ph, predicted}}(\mathbf{x})\right) \quad (2)$$

what is the $\mathbf{x}$ of the "most likely" mode.

$\Longrightarrow$ Minimization problem

Minimization Methods

Take the residual vector at a value $\mathbf{x}$:

$$\mathbf{R} = \frac{\text{computed} - \text{observed}}{\sigma} \tag{3}$$

Change as one varies $\mathbf{x}$:

$$\boldsymbol{\beta} = J^T\mathbf{R} \quad \text{where} \quad J = \frac{dR}{d\mathbf{x}} \tag{4}$$

J: Jacobi matrix.

Change parameters by

$$\Delta\mathbf{x} = -\alpha^{-1}\boldsymbol{\beta} \tag{5}$$

where the curvature matrix is

$$\alpha = J^T J \tag{6}$$

Then iterate until convergence.

Levenberg-Marquardt

Levenberg-Marquardt-method (LM-Method): Numerical method to minimize fit statistics

Levenberg, 1944, Q. Appl. Math 2, 164, Marquardt, 1963, SIAM J. Appl. Math. 11, 431

Modify simple gradient minimization by damping:
Replace α with α' where

$$\alpha'_{jj} = \begin{cases} \alpha_{jj}(1+\lambda) & \text{multiplicative damping} \\ \alpha_{jj}+\lambda & \text{additive damping} \end{cases} \tag{7}$$

i.e., the damped curvature matrix is (add. damping):

$$\alpha + \lambda \mathbb{1} \tag{8}$$

LM: adjust λ: Initially: want $\alpha \sim$ gradient, to lie in direction of steepest descent.

for additive damping: shrinks $\Delta p \Longrightarrow$ stablizing iteration by preconditioning the matrix α; in contrast, multiplicative damping will help the method against badly scaled problems

Levenberg-Marquardt

LM algorithm: decide how to change λ between steps:

Depending on how sum of squares (SOS) behaves at $\mathbf{x} + \Delta\mathbf{x}$:

- SOS decreased: $\lambda' = \lambda \cdot \mathrm{DROP}$
- SOS increased: drop p', set $\lambda' = \lambda \cdot \mathrm{BOOST}$

Note:

1. Assumes sum of squares-like likelihood

   ML won't work well on C-stat!!

2. Efficiency depends on values of BOOST and DROP.

Lampton (1997, Computers in Physics 11, 110): additive, DROP = 0.1, BOOST = 1.5 is well behaved, but not always

XSPEC does not allow changing these parameters, ISIS does