Journal of Theoretical Physics

Founded and Edited by M. Apostol **75** (2002)

ISSN 1453-4428

On the Levenberg-Marquardt minimization procedure

M. Apostol
Department of Theoretical Physics, Institute of Atomic Physics,
Magurele-Bucharest MG-6, POBox MG-35, Romania
email: apoma@theory.nipne.ro

Let $f(x)$ have a local minimum at x_m. We want to get it iteratively starting from x_0. We have the expansion

$$f(x) = f_0 + f_1(x - x_0) + \frac{1}{2}f_2(x - x_0)^2 \ , \tag{1}$$

where $f_0 = f(x_0)$, $f_1 = f'(x_0)$ and $f_2 = f''(x_0)$; this expansion reaches its minimum for $x_1$given by

$$f'(x_1) = f_1 + f_2(x_1 - x_0) = 0 \ , \tag{2}$$

i.e.

$$x_1 = x_0 - f_1/f_2 \ . \tag{3}$$

If $|x_1 - x_0| \ll |x_0|$, *i.e.* $|f_1| \ll |f_2 x_0|$, then

$$f(x_1) = f_0 - \frac{1}{2}f_1^2/f_2 \ , \tag{4}$$

and one can see that $f(x_1) < f_0$ providing $f_2 > 0$; therefore, x_1 determined as before takes us closer to the minimum value x_m. In general, if $f(x_1) < f_0$ we may iterate the procedure which leads toward the minimum position x_m.

If, on the contrary, $f(x_1) > f_0$, as, for instance for a very small $f_2 > 0$ where we may pass over the minimum, or for $f_2 < 0$ where we go towards a possible maximum value, then we must decrease the x-increment by taking

$$x_1 = x_0 - f_1/\lambda f_2 \ , \ \lambda \gg 1 \ , \tag{5}$$

such as $f(x_1) < f_0$; then we apply again the first step as given by (3).

The generalization to many variables is straightforward; the expansion (1) reads

$$f(...x_i, ...) = f_0 + \sum_i f_{1i}(x_i - x_i^0) + \frac{1}{2}\sum_{ij} f_{2ij}(x_i - x_i^0)(x_j - x_j^0) \ , \tag{6}$$

where $f_0 = f(...x_i^0...)$, $f_{1i} = \partial f/\partial x_i^0$ and $f_{2i} = \partial^2 f/\partial x_i^0 \partial x_j^0$, and the x_i-increments are solutions of the system

$$f_{1i} + \sum_j f_{2ij}(x_j^1 - x_j^0) = 0 \tag{7}$$

of linear equations. If $f(...x_i^1...) < f_0$ the procedure may be iterated; if, on the contrary, $f(...x_i^1...) > f_0$, then define new increments by

$$f_{1i} + \lambda f_{2ii}(x_i^1 - x_i^0) = 0 \ , \tag{8}$$

such as $f(...x_i^1...) < f_0$ and apply again the first step as given by (7). Equations (7) and (8) may be combined into one equation

$$f_{1i} + \sum_j f_{2ij}(1 + \lambda\delta_{ij})(x_j^1 - x_j^0) = 0 \ , \tag{9}$$

which for small λ amounts to (7), while for large λ goes over to (9). The procedure gives the local minima (either for f or for $-f$), and, of course, one needs a throughout investigation of the x-range in order to get all.[1]

References

[1] D. W. Marquardt, J. Ind. Soc. Appl. Math. **11** 431 (1963); see also W. H. Press et al, *Numerical Recipes*, Cambridge (1988).

© **J. Theor. Phys.** 2002, apoma@theor1.theory.nipne.ro