Quadrados mínimos não linear

(Método de Levenberg-Marquardt)

sistema não linear $\quad R(x) = 0$

quadrados mínimos $\quad \min_x \frac{1}{2}\|R(x)\|_2^2$.

$$f(x) = \frac{1}{2}\|R(x)\|_2^2 \, , \quad \nabla f(x) = J(x)^T R(x) \, ,$$

$$\nabla^2 f(x) = J(x)^T J(x) + S(x) \, , \quad \text{onde}$$

$$S(x) = \sum_{i=1}^{m} r_i(x) \nabla^2 r_i(x)$$ e

$$J(x) = \begin{bmatrix} \nabla r_1(x)^T \\ \vdots \\ \nabla r_m(x)^T \end{bmatrix}$$ $(m \times n)$ é a jacobiana de R.

<u>Newton</u>: resolver uma sequência de problemas de quadrados mínimos linear (descarta $S(x)$)

$$\min_d \frac{1}{2} \| J(x^k) d + R(x^k) \|_2^2 ,$$ e $x^{k+1} = x^k + d^k$.

Esses subproblemas aproximam o problema não linear ao redor de x^k ...

O método de Levenberg-Marquardt considera <u>regiões de confiança</u> em cada subproblema:

$$(*) \quad \min_{d} \tfrac{1}{2} \| J(x^k) d + R(x^k) \|_2^2$$

$$\text{s.a.} \quad \| d \|_2 \leq \Delta_k .$$

Note que este subproblema é globalmente equivalente à

$$\min_{d} \frac{1}{2} \| J(x^k) d + R(x^k) \|_2^2$$

$$\text{s.a.} \quad \frac{1}{2} \|d\|_2^2 \leq \frac{1}{2} \Delta_k^2 .$$

Este problema é convexo, e logo KKT é suficiente para caracterizar a solução global d^k:

KKT:

$$J(x^k)^T J(x^k)\, d^k + J(x^k)^T R(x^k) + \lambda d^k = 0,$$

$$\lambda \geq 0, \quad \lambda\left(\tfrac{1}{2}\|d^k\|_2^2 - \tfrac{1}{2}\Delta_k^2\right) = 0$$

$$\iff J(x^k)^T J(x^k)\, d^k + \lambda d^k = -J(x^k)^T R(x^k)$$

$$\lambda \geq 0, \quad \tfrac{1}{2}\lambda\left(\|d^k\|_2 - \Delta_k\right)\underbrace{\left(\|d^k\|_2 + \Delta_k\right)}_{>0} = 0$$

$$\Leftrightarrow \boxed{\begin{aligned} \left(J(x^k)^T J(x^k) + \lambda I\right) d^k &= -J(x^k)^T R(x^k) \\ \lambda \left(\|d^k\|_2 - \Delta_k\right) &= 0 \\ \lambda &\geq 0 \end{aligned}}$$

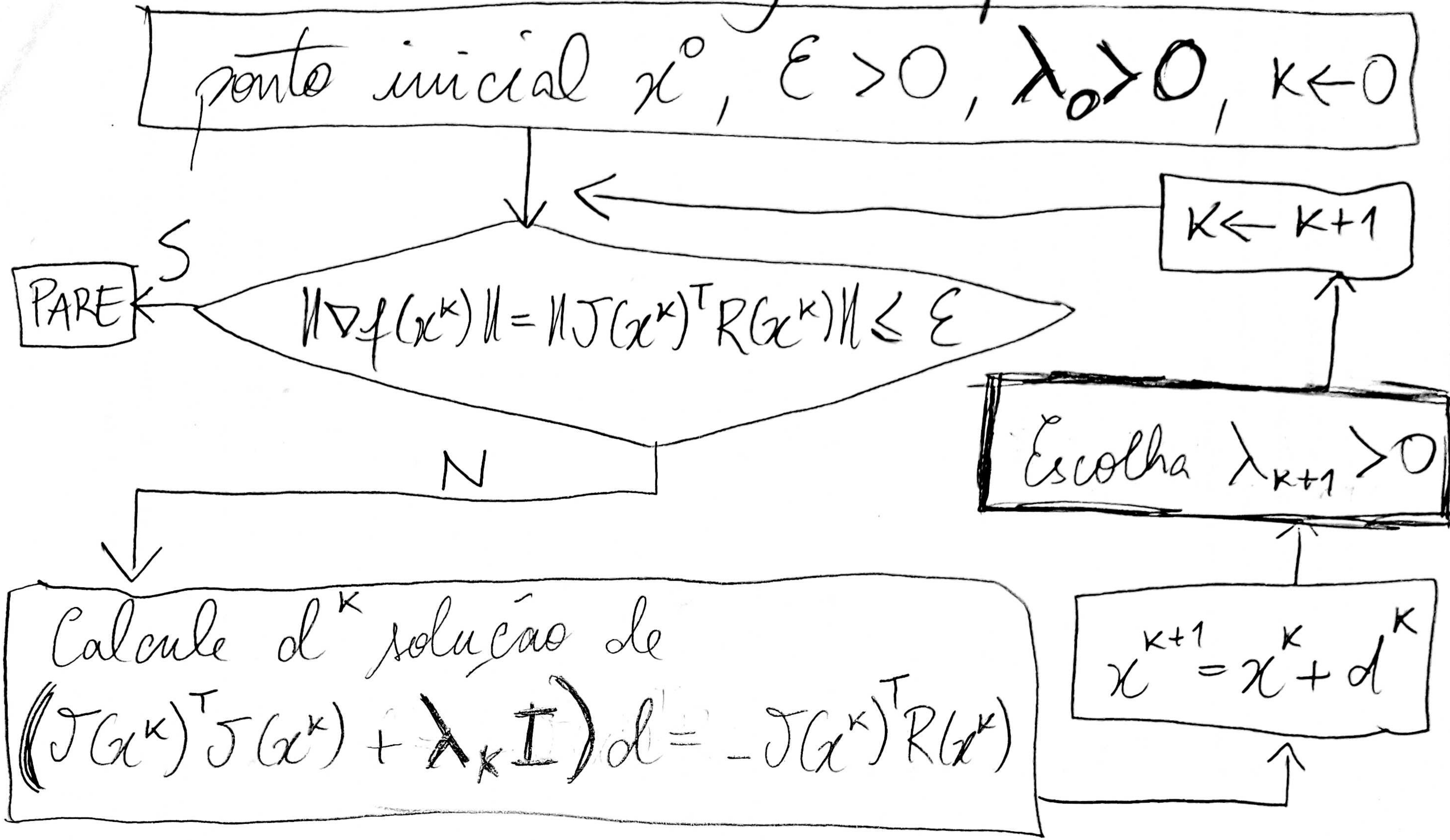
O termo λI $(\lambda > 0)$ pode ser visto como uma regularização de $J(x^k)^T J(x^k)$. De fato, $J(x^k)^T J(x^k) + \lambda I$ é definida positiva $\forall \lambda > 0$ (evita o problema de Gauss-Newton)

O método de Levenberg-Marquardt consiste em resolver uma sequência de "equações normais regularizadas"

$$\left(J(x^k)^T J(x^k) + \lambda_k I\right) d = -J(x^k)^T R(x^k)$$

em d e fazer $x^{k+1} = x^k + d^k$. A sequência $\lambda_k > 0$ é atualizada "heuristicamente". Observe que $\lambda_k = 0$ corresponde à Gauss-Newton.

Método de Levenberg-Marquardt



Observações:

1) Note que a condição de complementaridade $\lambda(\|d\|_2 - \Delta_k) = 0$ é descartada, assim como o próprio Δ_k.

2) Se $\lambda_k \gg 1$ então $J^t J + \lambda_k I \approx \lambda_k I$. Assim, d^k parecerá a direção de máxima descida.

3) Se $\lambda_k \approx 0$, então $d^k \approx$ Gauss-Newton.

4) É razoável que $\lambda_k \approx 0$ quando estamos próximos à uma solução de $R(x) \overset{!}{=} 0$. Possíveis escolhas:

$$\left.\begin{array}{l} \text{(i) } \lambda_k = \|R(x^k)\|_2^2 \\ \text{(ii) } \lambda_k = \|R(x^k)\|_2^\eta, \; \eta \in [1,2] \end{array}\right\} R \approx 0.$$

(iii) $\lambda_k = \|J(x^k)^T R(x^k)\| = \|\nabla f(x^k)\|$

($\approx$ estacionário/KKT).

10

5) Mostre que $(J(x)^T J(x) + \lambda I)\, d = -J(x)^T R(x)$ é a equação normal do problema de <u>quadrados mínimos linear</u>

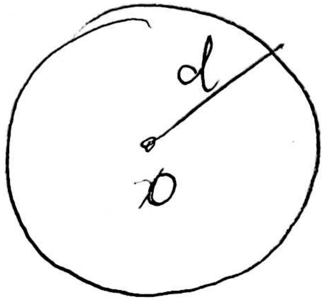
$$\min_{d} \frac{1}{2} \left\| \begin{bmatrix} J(x) \\ \sqrt{\lambda} I \end{bmatrix} d + \begin{bmatrix} R(x) \\ 0 \end{bmatrix} \right\|_2^2$$

Portanto pode-se aplicar técnicas anteriores (QR, Cholesky, SVD, GC).

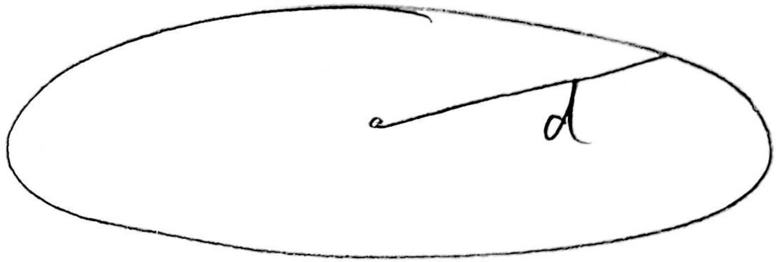
↳ Veja livro de Nocedal, seção 10.3.

6) Às vezes é conveniente mudar a região de confiança esférica $\|d\|_2 \leq \Delta_k$ para uma região elipsoidal

$$\|D_k d\|_2 \leq \Delta_k. \quad (D_k \succ 0 \text{ diagonal})$$



$\|d\|_2 \leq \Delta_k$



$\|D_k d\|_2 \leq \Delta_k.$

Mostre que a equação correspondente é:

$$\left(J(x^k)^T J(x^k) + \lambda_k D_k^2 \right) d = -J(x^k)^T R(x^k),$$

e que o problema de quadrados mínimos linear associado é

$$\min_d \frac{1}{2} \left\| \begin{bmatrix} J(x^k) \\ \sqrt{\lambda_k} D_k \end{bmatrix} d + \begin{bmatrix} R(x^k) \\ 0 \end{bmatrix} \right\|_2^2$$