# Non linear Least Squares

## Lectures for PHD course on Numerical Optimization

Enrico Bertolazzi

DII – Università di Trento



## Notes

# Outline





## Notes

- An important class on minimization problem when $\mathsf{f} : \mathbb{R}^n \mapsto \mathbb{R}$ is the nonlinear least squares and takes the form:

$$\mathsf{f}(\boldsymbol{x}) = \frac{1}{2}\sum_{i=1}^{m} F_i(\boldsymbol{x})^2, \qquad m \geq n$$

- When $n = m$ finding the minimum coincide to finding the solution of the non linear system $\mathbf{F}(\boldsymbol{x}) = \mathbf{0}$ where:

$$\mathbf{F}(\boldsymbol{x}) = \big(F_1(\boldsymbol{x}), F_2(\boldsymbol{x}), \ldots, F_n(\boldsymbol{x})\big)^T$$

- Thus, special methods developed for the solution of nonlinear least squares can be used for the solution of nonlinear systems, but not the converse if $m > n$.



## Notes

## Example

Consider the the following fitting model

$$M(\boldsymbol{x}, t) = x_3 \exp(x_1 t) + x_4 \exp(x_3 t)$$

which can be used to fit some data. The model depend on the parameters $\boldsymbol{x} = (x_1, x_2, x_3, x_4)^T$. If we have a number of points

$$(t_k, y_k)^T, \qquad k = 1, 2, \ldots, m$$

we want to find the parameters $\boldsymbol{x}$ such that $\frac{1}{2}\sum_{k=1}^{m}(M(\boldsymbol{x}, t_k) - y_k)^2$ is minimum. Defining

$$F_k(\boldsymbol{x}) = M(\boldsymbol{x}, t_k) - y_k, \qquad k = 1, 2, \ldots, m$$

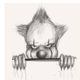
then can be viewed as a non linear least squares problem.



Notes

- To solve nonlinear least squares problem, we can use any of the previously discussed method. For example BFGS or Newton method with globalization techniques.
- If for example we use Newton method we need to compute

$$
\begin{aligned}
\nabla^2 \mathsf{f}(\boldsymbol{x}) &= \nabla^2 \frac{1}{2}\sum_{i=1}^{m} F_i(\boldsymbol{x})^2 = \frac{1}{2}\sum_{i=1}^{m} \nabla^2 F_i(\boldsymbol{x})^2 \\
&= \frac{1}{2}\sum_{i=1}^{m} \nabla(2F_i(\boldsymbol{x})\nabla F_i(\boldsymbol{x}))^T \\
&= \sum_{i=1}^{m} \nabla F_i(\boldsymbol{x})^T \nabla F_i(\boldsymbol{x}) + \sum_{i=1}^{m} F_i(\boldsymbol{x})\nabla^2 F_i(\boldsymbol{x})
\end{aligned}
$$





Notes

- If we define

$$
\boldsymbol{J}(\boldsymbol{x}) = \begin{pmatrix} \nabla F_1(\boldsymbol{x}) \\ \nabla F_2(\boldsymbol{x}) \\ \vdots \\ \nabla F_m(\boldsymbol{x}) \end{pmatrix}
$$

  then we can write

$$
\nabla^2 \mathsf{f}(\boldsymbol{x}) = \boldsymbol{J}(\boldsymbol{x})^T \boldsymbol{J}(\boldsymbol{x}) + \sum_{i=1}^{m} F_i(\boldsymbol{x}) \nabla^2 F_i(\boldsymbol{x})
$$

- However, in practical problem normally $\boldsymbol{J}(\boldsymbol{x})$ is known, while $\nabla^2 F_i(\boldsymbol{x})$ is not known or impractical to compute.

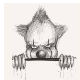


Notes

- A common approximation is given by neglecting the terms $\nabla^2 F_i(\boldsymbol{x})$ obtaining,

$$\nabla^2 \mathsf{f}(\boldsymbol{x}) \approx \boldsymbol{J}(\boldsymbol{x})^T \boldsymbol{J}(\boldsymbol{x})$$

- This choice can be appropriate near the solution if $n = m$ in solving nonlinear system. In fact near the solution we have $F_i(\boldsymbol{x}) \approx 0$ so that the contribution of the neglected term is small.
- This choice is not good when near the minimum we have large residual (i.e. $\|\mathbf{F}(\boldsymbol{x})\|$ is large) because the contribution of $\nabla^2 F_i(\boldsymbol{x})$ cant be neglected.





Notes

# Outline





## Notes

# References


J. Stoer and R. Bulirsch
Introduction to numerical analysis
Springer-Verlag, Texts in Applied Mathematics, **12**, 2002.

J. E. Dennis, Jr. and Robert B. Schnabel
Numerical Methods for Unconstrained Optimization and Nonlinear Equations
SIAM, Classics in Applied Mathematics, **16**, 1996.




## Notes