System of nonlinear equations with separable variables

Let $\mathbf{f} : \mathsf{R}_n \to \mathsf{R}_n$ is a function of variable $\mathbf{x} \in \mathsf{R}_n$. Then, if $\boldsymbol{\alpha}$ is a solution of the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$, the function $\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2 = \frac{1}{2}\mathbf{f}^{\mathrm{T}}(\mathbf{x})\mathbf{f}(\mathbf{x})$ has global minimum at the point $\boldsymbol{\alpha}$.

Let $\mathbf{f} : \mathsf{R}_n \to \mathsf{R}_m$, $m > n$, is a function of variable $\mathbf{x} \in \mathsf{R}_n$. Then, if the function $\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2$ has global minimum at the point $\boldsymbol{\alpha}$, we say, that $\boldsymbol{\alpha}$ is a solution of the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$ according to the criterion of least sum of squares of residuals.

Let $\mathbf{f} : \mathsf{R}_n \to \mathsf{R}_m$, $m \geq n$, is a function of variable $\mathbf{x} \in \mathsf{R}_n$ such that

$$\mathbf{f}(\mathbf{x}) \equiv \mathbf{f}\begin{pmatrix} \mathbf{x}_1 \\ \mathbf{x}_2 \end{pmatrix} = \mathbf{b}(\mathbf{x}_2) - \mathbf{A}(\mathbf{x}_2)\mathbf{x}_1 \ ,$$

$\mathbf{x}_1 \in \mathsf{R}_k$, $\mathbf{x}_2 \in \mathsf{R}_{n-k}$, $\mathbf{A} : \mathsf{R}_{n-k} \to \mathsf{R}_{m\times k}$, $\mathbf{b} : \mathsf{R}_{n-k} \to \mathsf{R}_m$ (the variables of the function $\mathbf{f}(\mathbf{x})$ are *separable*, i.e. $\mathbf{f}(\mathbf{x})$ is a linear function of $\mathbf{x}_1$ and nonlinear function of $\mathbf{x}_2$). Let $\mathbf{f}_2 : \mathsf{R}_{n-k} \to \mathsf{R}_m$ is a function of $\mathbf{x}_2$ such that¶

$$\mathbf{f}_2(\mathbf{x}_2) = \mathbf{b}(\mathbf{x}_2) - \mathbf{A}(\mathbf{x}_2)\mathbf{A}^+(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2) = \left[\mathbf{1}_m - \mathbf{A}(\mathbf{x}_2)\mathbf{A}^+(\mathbf{x}_2)\right]\mathbf{b}(\mathbf{x}_2) \ .$$

Then[1]

$$\mathbf{f}(\boldsymbol{\alpha}) \equiv \mathbf{f}\begin{pmatrix} \boldsymbol{\alpha}_1 \\ \boldsymbol{\alpha}_2 \end{pmatrix} = \mathbf{0} \ \Leftrightarrow \ \mathbf{f}_2(\boldsymbol{\alpha}_2) = \mathbf{0} \ \wedge \ \boldsymbol{\alpha}_1 = \mathbf{A}^+(\boldsymbol{\alpha}_2)\mathbf{b}(\boldsymbol{\alpha}_2) \ .$$

In words, if $\boldsymbol{\alpha}$ is a solution of the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$†, $\boldsymbol{\alpha}_2$ is a solution of the equation $\mathbf{f}_2(\boldsymbol{\alpha}_2) = \mathbf{0}$ and $\boldsymbol{\alpha}_1 = \mathbf{A}^+(\boldsymbol{\alpha}_2)\mathbf{b}(\boldsymbol{\alpha}_2)$. And also conversely (which is much more valuable from the practical point of view), if $\boldsymbol{\alpha}_2$ is a solution of the equation $\mathbf{f}_2(\boldsymbol{\alpha}_2) = \mathbf{0}$ and $\boldsymbol{\alpha}_1 = \mathbf{A}^+(\boldsymbol{\alpha}_2)\mathbf{b}(\boldsymbol{\alpha}_2)$, $\boldsymbol{\alpha}$ is a solution of the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$§.

This theorem says, that if the variables of the function $\mathbf{f}(\mathbf{x})$ are separable, we can solve the equation $\mathbf{f}_2(\boldsymbol{\alpha}_2) = \mathbf{0}_{n-k}$ instead of the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}_n$ and compute $\boldsymbol{\alpha}_1$ "later".

Let us suppose, that we use a derivative algorithm to solve the equation, which requires the knowledge of the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2) \sim (m, n-k)$ of the function $\mathbf{f}_2(\mathbf{x}_2)$‡

$$\mathbf{J}_2(\mathbf{x}_2) = \left\{\frac{\partial f_{2i}}{\partial x_{2j}}\right\} \quad , \quad i = 1, \ldots, m \quad , \quad j = 1, \ldots, n-k$$

and also the gradient (i.e. the vector of the first derivatives) of the function $\frac{1}{2}\|\mathbf{f}_2(\mathbf{x}_2)\|^2$

$$grad\left[\tfrac{1}{2}\mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)\right] = \mathbf{J}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2) \ .$$

However, usually we know nothing but the Jacobi matrix $\mathbf{J}(\mathbf{x}) \sim (m, n)$ of the function $\mathbf{f}(\mathbf{x})$

$$\mathbf{J}(\mathbf{x}) = \left\{\frac{\partial f_i}{\partial x_j}\right\} \quad , \quad i = 1, \ldots, m \quad , \quad j = 1, \ldots, n \ ,$$

whose "left" submatrix of type (m, k) is the matrix $\mathbf{A}(\mathbf{x}_2)$. The composition of the matrix $\mathbf{J}(\mathbf{x})$ and vector $\mathbf{x}_1 = \mathbf{A}^+(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2)$ leads to the matrix $(\,\mathbf{A}(\mathbf{x}_2) \quad \mathbf{J}_{n-k}(\mathbf{x}_2)\,)$.

¶ If $\mathbf{f}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{x}_1 = \mathbf{0}$, then $\mathbf{x}_1 = \mathbf{A}^+\mathbf{b}$. If we introduce this solution "back" into the relation for $\mathbf{f}(\mathbf{x})$, we get $\mathbf{f}(\mathbf{x}) = \mathbf{b} - \mathbf{A}\mathbf{A}^+\mathbf{b} = (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{b}$. In different words, the function $\mathbf{f}_2(\mathbf{x}_2)$ is a composition of the function $\mathbf{f}(\mathbf{x})$ and vector $\mathbf{x}_1 = \mathbf{A}^+(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2)$.

† according to the criterion of least sum of squares of residuals

§ If the function $\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2$ has global minimum at the point $\boldsymbol{\alpha}$, the function $\frac{1}{2}\|\mathbf{f}_2(\mathbf{x}_2)\|^2$ has global minimum at the point $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_1 = \mathbf{A}^+(\boldsymbol{\alpha}_2)\mathbf{b}(\boldsymbol{\alpha}_2)$. If the function $\frac{1}{2}\|\mathbf{f}_2(\mathbf{x}_2)\|^2$ has global minimum at the point $\boldsymbol{\alpha}_2$ and $\boldsymbol{\alpha}_1 = \mathbf{A}^+(\boldsymbol{\alpha}_2)\mathbf{b}(\boldsymbol{\alpha}_2)$, the function $\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2$ has global minimum at the point $\boldsymbol{\alpha}$.

‡ and that both the matrix $\mathbf{A}(\mathbf{x}_2)$ and the vector $\mathbf{b}(\mathbf{x}_2)$ are continuously differentiable at the point $\mathbf{x}_2$ and the matrix $\mathbf{A}(\mathbf{x}_2)$ has continuous (i.e. actually constant, but not necessarily full) rank $r(\mathbf{A}) \leq n$ at the point $\mathbf{x}_2$, thanks to which the matrix $\mathbf{A}^+(\mathbf{x}_2)$ is continuously differentiable at the point $\mathbf{x}_2$, too

In the following paragraph we will build a mathematical formalism, with the aid of which we will express the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2)$ analytically as a function of the matrix $(\mathbf{A}(\mathbf{x}_2)\ \ \mathbf{J}_{n-k}(\mathbf{x}_2))$. However, first we will extend the validity of the multiplication and transposition operation of matrices to the "three-dimensional" case. Thus, let the multiplication of a three-dimensional and two-dimensional matrix takes its course in the first and second dimension, while the result is **always** a three-dimensional matrix (the second dimension can be equal to one). Analogically, let the transposition of a three-dimensional matrix consists in the exchange of the first and second dimension (the first and second index of **all** elements)¶.

Let $\mathbf{f} : \mathsf{R}_p \to \mathsf{R}_{m\times n}$ is a function of variable $\mathbf{x} \in \mathsf{R}_p$ and $\mathcal{D} \equiv \mathcal{D}_{\mathbf{x}\in\mathsf{R}_p} : \mathsf{R}_{m\times n} \to \mathsf{R}_{m\times n\times p}$ is a projection such that

$$\mathcal{D}_{\mathbf{x}\in\mathsf{R}_p}\left[\mathbf{f}(\mathbf{x})\right] = \left\{\frac{\partial f_{ij}}{\partial x_k}\right\} \quad , \quad i=1,\ldots,m \quad , \quad j=1,\ldots,n \quad , \quad k=1,\ldots,p \; .$$

Let $\mathbf{A},\mathbf{B} : \mathsf{R}_p \to \mathsf{R}_{m\times n}$, $\mathbf{a},\mathbf{b} : \mathsf{R}_p \to \mathsf{R}_{m\times 1}$ (i.e. $\mathbf{a}^{\mathrm{T}},\mathbf{b}^{\mathrm{T}} : \mathsf{R}_p \to \mathsf{R}_{1\times m}$), $a,b : \mathsf{R}_p \to \mathsf{R}_{1\times 1}$ are functions of variable $\mathbf{x} \in \mathsf{R}_p$. Then

$$\mathcal{D}(\mathbf{AB}) = \mathcal{D}(\mathbf{A})\mathbf{B} + \mathbf{A}\mathcal{D}(\mathbf{B}) \; ,$$

$$\mathcal{D}(\mathbf{Ab}) = \mathcal{D}(\mathbf{A})\mathbf{b} + \mathbf{A}\mathcal{D}(\mathbf{b}) \; ,$$
$$\mathcal{D}(\mathbf{a}^{\mathrm{T}}\mathbf{b}) = \mathcal{D}\left(\mathbf{a}^{\mathrm{T}}\right)\mathbf{a} + \mathbf{a}^{\mathrm{T}}\mathcal{D}(\mathbf{b}) \; ,$$
$$\mathcal{D}(a\mathbf{B}) = \mathcal{D}(a)\mathbf{B} + a\mathcal{D}(\mathbf{B}) \; ,$$
$$\mathcal{D}(a\mathbf{b}) = \mathcal{D}(a)\mathbf{b} + a\mathcal{D}(\mathbf{b}) \; ,$$
$$\mathcal{D}(ab) = \mathcal{D}(a)b + a\mathcal{D}(b) \; ,$$
$$\mathcal{D}\left(\mathbf{A}^{\mathrm{T}}\right) = \left[\mathcal{D}(\mathbf{A})\right]^{\mathrm{T}} \; ,$$
$$\mathcal{D}\left(\mathbf{a}^{\mathrm{T}}\right) = \left[\mathcal{D}(\mathbf{a})\right]^{\mathrm{T}} \; .$$

It is necessary to take into consideration that while the multiplication and transposition of matrices takes its course in the "paper plane", the differentiation with respect to the elements of a vector goes in the third dimension, i.e. "perpendicularly to the paper". Thus, the Jacobi matrix $\mathbf{J}(\mathbf{x}) = \mathcal{D}_{\mathbf{x}\in\mathsf{R}_n}\left[\mathbf{f}(\mathbf{x})\right] = \left(\mathcal{D}_{\mathbf{x}_1\in\mathsf{R}_k}\left[\mathbf{f}(\mathbf{x})\right]\ \ \mathcal{D}_{\mathbf{x}_2\in\mathsf{R}_{n-k}}\left[\mathbf{f}(\mathbf{x})\right]\right) = \left(\mathbf{A}(\mathbf{x}_2)\ \ \mathcal{D}_{\mathbf{x}_2\in\mathsf{R}_{n-k}}\left[\mathbf{f}(\mathbf{x})\right]\right)$ is a matrix of type $(m,1,n)$, the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2) = \mathcal{D}_{\mathbf{x}_2\in\mathsf{R}_{n-k}}\left[\mathbf{f}_2(\mathbf{x}_2)\right]$ is a matrix of type $(m,1,n-k)$ and the gradient of the function $\frac{1}{2}\mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)$ is a vector of type $(1,1,n-k)$ in this and several following paragraphs. Obviously

$$\mathcal{D}_{\mathbf{x}_2\in\mathsf{R}_{n-k}}\left[\mathbf{f}(\mathbf{x})\right] = \mathcal{D}(\mathbf{b}) - \mathcal{D}(\mathbf{A})\mathbf{x}_1 \; .$$

The composition of this matrix† and vector $\mathbf{x}_1 = \mathbf{A}^{+}(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2)$ leads to the matrix $\mathbf{J}_{n-k}(\mathbf{x}_2)$. Then§

$$\mathbf{J}_{n-k}(\mathbf{x}_2) = \mathcal{D}(\mathbf{b}) - \mathcal{D}(\mathbf{A})\mathbf{A}^{+}\mathbf{b} \; ,$$

$$\begin{aligned}\mathbf{J}_2(\mathbf{x}_2) &= \mathcal{D}\left[\mathbf{f}_2(\mathbf{x}_2)\right] = \mathcal{D}\left[(\mathbf{1}_m - \mathbf{A}\mathbf{A}^{+})\mathbf{b}\right] = (\mathbf{1}_m - \mathbf{A}\mathbf{A}^{+})\mathcal{D}(\mathbf{b}) - \left[\mathcal{D}(\mathbf{A})\mathbf{A}^{+} + \mathbf{A}\mathcal{D}(\mathbf{A}^{+})\right]\mathbf{b} = \\ &= \mathcal{D}(\mathbf{b}) - \mathcal{D}(\mathbf{A})\mathbf{A}^{+}\mathbf{b} - \mathbf{A}\mathbf{A}^{+}\mathcal{D}(\mathbf{b}) - \mathbf{A}\mathcal{D}(\mathbf{A}^{+})\mathbf{b} = \mathbf{J}_{n-k}(\mathbf{x}_2) - \mathbf{A}\left[\mathbf{A}^{+}\mathcal{D}(\mathbf{b}) + \mathcal{D}(\mathbf{A}^{+})\mathbf{b}\right] \; .\end{aligned}$$

¶ $\mathbf{AB} = \mathbf{C}$, $\mathbf{Ab} = \mathbf{c}$, $\mathbf{A} \sim (m,n,p)$, $\mathbf{B} \sim (n,k)$, $\mathbf{C} \sim (m,k,p)$, $\mathbf{b} \sim (n,1)$, $\mathbf{c} \sim (m,1,p)$, $\mathbf{C}^{\mathrm{T}} \sim (k,m,p)$, $\mathbf{c}^{\mathrm{T}} \sim (1,m,p)$. Thus, in the first example ($\mathbf{AB} = \mathbf{C}$) we multiply sequentially p matrices of type (m,n) with one and the same matrix $\mathbf{B}$ of type (n,k) obtaining p matrices of type (m,k).

† I will say once more, that it is the "right" submatrix of type $(m,1,n-k)$ of the Jacobi matrix $\mathbf{J}(\mathbf{x})$

§ hitherto we had to distinguish between $\mathcal{D}_{\mathbf{x}\in\mathsf{R}_n}$, $\mathcal{D}_{\mathbf{x}_1\in\mathsf{R}_k}$ and $\mathcal{D}_{\mathbf{x}_2\in\mathsf{R}_{n-k}}$, starting from this point we solely differentiate with respect to $\mathbf{x}_2 \in \mathsf{R}_{n-k}$

From the definition of the Moore-Penrose pseudoinverse matrix¶ if follows that

$$\mathbf{A}^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2) = \mathbf{A}^{\mathrm{T}}(\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{b} = \left[\mathbf{A}^{\mathrm{T}} - \mathbf{A}^{\mathrm{T}}\left(\mathbf{A}\mathbf{A}^+\right)^{\mathrm{T}}\right]\mathbf{b} = \left[\mathbf{A}^{\mathrm{T}} - \left(\mathbf{A}\mathbf{A}^+\mathbf{A}\right)^{\mathrm{T}}\right]\mathbf{b} = \mathbf{0}$$

and the gradient of the function $\frac{1}{2}\mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)$ can easily be obtained by mere "replacement" of the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2)$ with the already well-known matrix $\mathbf{J}_{n-k}(\mathbf{x}_2)$

$$grad\left[\tfrac{1}{2}\mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)\right] = \mathbf{J}_{n-k}^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2) - \left[\mathbf{A}^+\mathcal{D}(\mathbf{b}) + \mathcal{D}(\mathbf{A}^+)\mathbf{b}\right]^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2) = \mathbf{J}_{n-k}^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)\ .$$

No relation between the matrices $\mathbf{J}_2(\mathbf{x}_2)$ and $\mathbf{J}_{n-k}(\mathbf{x}_2)$ follows from this fact yet. However, it is clear that both the shift vector $\boldsymbol{\delta}_2 = -\mathbf{J}_2^+(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)$ (if $m = n$, we talk about the Newton-Raphson method, if $m > n$, we talk about the Gauss-Newton method) and the shift vector $\boldsymbol{\delta}_2 = -\mathbf{J}_{n-k}^+(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)$ are descent vectors of the function $\frac{1}{2}\|\mathbf{f}_2(\mathbf{x}_2)\|^2$ at the point $\mathbf{x}_2 \neq \boldsymbol{\alpha}_2$, i.e.

$$\left\{grad\left[\tfrac{1}{2}\mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2)\right]\right\}^{\mathrm{T}}\boldsymbol{\delta}_2 < 0\ ,$$

if the matrices $\mathbf{J}_2(\mathbf{x}_2)$ and $\mathbf{J}_{n-k}(\mathbf{x}_2)$ have full rank† at the point $\mathbf{x}_2$. Thus, if we substitute the matrix $\mathbf{J}_{n-k}(\mathbf{x}_2)$ for the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2)$ in the Newton-Raphson or Gauss-Newton method, both the methods will still converge to the right solution. But usually much more slowly (even in comparison with methods, which adopt the shift vector $\boldsymbol{\delta} = -\mathbf{J}^+(\mathbf{x})\mathbf{f}(\mathbf{x})$ to solve the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$, if we calculate the initial estimate of $\boldsymbol{\alpha}_1$ from the initial estimate of $\boldsymbol{\alpha}_2$ according to the relation $\mathbf{x}_1 = \mathbf{A}^+(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2)$), thanks to which the described simplification is almost unusable in practice§.

The relation between the matrices $\mathbf{J}_2(\mathbf{x}_2)$ and $\mathbf{J}_{n-k}(\mathbf{x}_2)$ can be derived even without an explicit expression of the three-dimensional matrix $\mathcal{D}(\mathbf{A}^+)$ of type $(k, m, n-k)$. It holds that

$$\mathbf{J}_2(\mathbf{x}_2) = \mathcal{D}\left[\mathbf{f}_2(\mathbf{x}_2)\right] = \mathcal{D}\left[(\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{b}\right] = -\mathcal{D}(\mathbf{A}\mathbf{A}^+)\mathbf{b} + (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{b})$$

and from the definition of the Moore-Penrose pseudoinverse matrix¶ it follows that[1]

$$\mathcal{D}(\mathbf{A}\mathbf{A}^+\mathbf{A}) = \mathcal{D}(\mathbf{A}) = \mathcal{D}(\mathbf{A}\mathbf{A}^+)\mathbf{A} + \mathbf{A}\mathbf{A}^+\mathcal{D}(\mathbf{A})\ ,$$

$$\mathcal{D}(\mathbf{A}\mathbf{A}^+)\mathbf{A}\mathbf{A}^+ = (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{A})\mathbf{A}^+\ ,$$

$$\mathcal{D}(\mathbf{A}\mathbf{A}^+\mathbf{A}\mathbf{A}^+) = \mathcal{D}(\mathbf{A}\mathbf{A}^+) = \mathcal{D}(\mathbf{A}\mathbf{A}^+)\mathbf{A}\mathbf{A}^+ + \mathbf{A}\mathbf{A}^+\mathcal{D}(\mathbf{A}\mathbf{A}^+)\ ,$$

$$\mathbf{A}\mathbf{A}^+\mathcal{D}(\mathbf{A}\mathbf{A}^+) = \left\{\left[\mathcal{D}(\mathbf{A}\mathbf{A}^+)\right]^{\mathrm{T}}(\mathbf{A}\mathbf{A}^+)^{\mathrm{T}}\right\}^{\mathrm{T}} = \left\{\mathcal{D}\left[(\mathbf{A}\mathbf{A}^+)^{\mathrm{T}}\right]\mathbf{A}\mathbf{A}^+\right\}^{\mathrm{T}} = \left[\mathcal{D}(\mathbf{A}\mathbf{A}^+)\mathbf{A}\mathbf{A}^+\right]^{\mathrm{T}}\ ,$$

$$\mathcal{D}(\mathbf{A}\mathbf{A}^+) = (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{A})\mathbf{A}^+ + \left[(\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{A})\mathbf{A}^+\right]^{\mathrm{T}}\ ,$$

$$\begin{aligned}\mathbf{J}_2(\mathbf{x}_2) &= -(\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{A})\mathbf{A}^+\mathbf{b} - \left[\mathcal{D}(\mathbf{A})\mathbf{A}^+\right]^{\mathrm{T}}(\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{b} + (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathcal{D}(\mathbf{b}) = \\ &= (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{J}_{n-k}(\mathbf{x}_2) - [\mathcal{D}(\mathbf{A})\mathbf{A}^+]^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2)\ .\end{aligned}$$

¶

$$\mathbf{A}\mathbf{A}^+\mathbf{A} = \mathbf{A}\ ,\quad \mathbf{A}^+\mathbf{A}\mathbf{A}^+ = \mathbf{A}^+\ ,$$

$$(\mathbf{A}\mathbf{A}^+)^{\mathrm{T}} = \mathbf{A}\mathbf{A}^+\ ,\quad (\mathbf{A}^+\mathbf{A})^{\mathrm{T}} = \mathbf{A}^+\mathbf{A}$$

† quite generally, quadratic form

$$\mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{A}^+\mathbf{b} = \mathbf{b}^{\mathrm{T}}\mathbf{A}\mathbf{A}^+\mathbf{A}\mathbf{A}^+\mathbf{b} = \mathbf{b}^{\mathrm{T}}\left(\mathbf{A}\mathbf{A}^+\right)^{\mathrm{T}}\mathbf{A}\mathbf{A}^+\mathbf{b} = \left(\mathbf{A}^+\mathbf{b}\right)^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{A}\mathbf{A}^+\mathbf{b}\ ,$$

$\mathbf{A} \sim (m,n)$, $\mathbf{b} \sim (m,1)\ \wedge\ \mathbf{b} \neq \mathbf{0}$, $m \geq n$, is positive if and only if the matrix $\mathbf{A}^{\mathrm{T}}\mathbf{A}$ is positive definite and the vector $\mathbf{A}^+\mathbf{b}$ is nonzero, i.e. if the matrix $\mathbf{A}$ has full rank

§ especially close to a local minimum the number of iterations is unbearably large

It is instantly evident that at least near the global minimum, where the l_2-norm of the vector $\mathbf{f}_2(\mathbf{x}_2)$ is "small", the importance of the second term decreases¶. It is just necessary to remember the commonly accepted approximation of the Hess matrix $\mathbf{H}(\mathbf{x}) \sim (n,n)$ of the function $\frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2$

$$\mathbf{H}(\mathbf{x}) = \left\{ \frac{\partial^2 \frac{1}{2}\|\mathbf{f}(\mathbf{x})\|^2}{\partial x_i \partial x_j} \right\} \quad , \quad i = 1,\ldots,n \quad , \quad j = 1,\ldots,n \ ,$$

$$\mathbf{H}(\mathbf{x}) = \mathbf{J}^{\mathrm{T}}(\mathbf{x})\mathbf{J}(\mathbf{x}) + \mathbf{C}(\mathbf{x})\mathbf{f}(\mathbf{x}) \ ,$$

$$\mathbf{C}(\mathbf{x}) = \left\{ \frac{\partial^2 f_i}{\partial x_j \partial x_k} \right\} \quad , \quad i = 1,\ldots,m \quad , \quad j = 1,\ldots,n \quad , \quad k = 1,\ldots,n \ ,$$

$$\mathbf{H}(\mathbf{x}) \approx \mathbf{J}^{\mathrm{T}}(\mathbf{x})\mathbf{J}(\mathbf{x}) \ ,$$

which forms the basis of the Gauss-Newton method of solving the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}$. Hence, let us hope that the second (unpleasantly complicated) term $-[\mathcal{D}(\mathbf{A})\mathbf{A}^+]^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2)$ in the relation for $\mathbf{J}_2(\mathbf{x}_2)$ can be neglected without serious consequences, i.e. just the first term is enough to be taken into account[2]

$$\mathbf{J}_2(\mathbf{x}_2) \approx (\mathbf{1}_m - \mathbf{A}\mathbf{A}^+)\mathbf{J}_{n-k}(\mathbf{x}_2)$$

and direct manipulation with the three-dimensional matrix $\mathcal{D}(\mathbf{A})$ of type $(m,k,n-k)$ can be avoided (while both computer time and memory are saved). However, it should be noted that all the necessary derivatives have to be computed anyway, since they are "implicitly" included in the matrix $\mathbf{J}_{n-k}(\mathbf{x}_2) = \mathcal{D}(\mathbf{b}) - \mathcal{D}(\mathbf{A})\mathbf{A}^+\mathbf{b}$†.

Thus, the described approximation is valuable especially for those, who want to easily and quickly modify almost any derivative algorithm of solving the equation $\mathbf{f}(\boldsymbol{\alpha}) = \mathbf{0}_n$ to a (possibly) more efficient algorithm of solving the "equivalent" equation $\mathbf{f}_2(\boldsymbol{\alpha}_2) = \mathbf{0}_{n-k}$ and are not willing to get over the loss of the already derived relations for calculation of the necessary partial derivatives. Since if we differentiate numerically, we only manage with $n-k+1$ values (at different points) of the sum of squares of residuals $\|\mathbf{f}_2(\mathbf{x}_2)\|^2$ in each iteration (to compute the gradient of the function $\frac{1}{2}\|\mathbf{f}_2(\mathbf{x}_2)\|^2$) and $n-k+1$ values of the function $\mathbf{f}_2(\mathbf{x}_2)$ (to compute the Jacobi matrix $\mathbf{J}_2(\mathbf{x}_2)$). The algorithm of calculation of $\|\mathbf{f}_2(\mathbf{x}_2)\|^2$ becomes trivial after performing a suitable decomposition§ of the matrix $\mathbf{A}(\mathbf{x}_2)$, the algorithm of calculation of $\mathbf{f}_2(\mathbf{x}_2)$ is just a little more complicated. The question naturally arises if it is not more profitable to resign the use of analytical derivatives to the use of derivatives calculated according to common difference formulae.

If $r(\mathbf{A}) = k$ (i.e. if the matrix $\mathbf{A}(\mathbf{x}_2)$ has full rank), then

$$\mathbf{A} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{U} \\ \mathbf{0} \end{pmatrix} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_k \\ \mathbf{0} \end{pmatrix} \mathbf{U} = \mathbf{Q}_k^{\mathrm{T}}\mathbf{U} \ ,$$

$$\mathbf{A}^+ = \mathbf{U}^{-1}\mathbf{Q}_k \ ,$$

$$\mathbf{f}_2(\mathbf{x}_2) = (\mathbf{1}_m - \mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k)\mathbf{b} \ ,$$

$$\|\mathbf{f}_2(\mathbf{x}_2)\|^2 = \|\mathbf{Q}_{m-k}\mathbf{b}\|^2 \ \ddagger \ ,$$

¶ However, the sum of squares of residuals can be large even at the global minimum if $m > n$.

† I will say once more, that the matrix $\mathbf{J}_{n-k}(\mathbf{x}_2)$ is composed of the "right" submatrix of type $(m,1,n-k)$ of the Jacobi matrix $\mathbf{J}(\mathbf{x})$ and vector $\mathbf{x}_1 = \mathbf{A}^+(\mathbf{x}_2)\mathbf{b}(\mathbf{x}_2)$

§ we only describe the QR- and singular value decomposition, which are the commonest

‡

$$\mathbf{f}_2(\mathbf{x}_2) = \left[\mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{m-k} \end{pmatrix} \mathbf{Q} - \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_k & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{Q}\right] \mathbf{b} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{1}_{m-k} \end{pmatrix} \mathbf{Q}\mathbf{b} = \mathbf{Q}_{m-k}^{\mathrm{T}}\mathbf{Q}_{m-k}\mathbf{b} \ ,$$

$$\|\mathbf{f}_2(\mathbf{x}_2)\|^2 = \mathbf{f}_2^{\mathrm{T}}(\mathbf{x}_2)\mathbf{f}_2(\mathbf{x}_2) = \mathbf{b}^{\mathrm{T}}\mathbf{Q}_{m-k}^{\mathrm{T}}\mathbf{Q}_{m-k}\mathbf{Q}_{m-k}^{\mathrm{T}}\mathbf{Q}_{m-k}\mathbf{b} = \mathbf{b}^{\mathrm{T}}\mathbf{Q}_{m-k}^{\mathrm{T}}\mathbf{Q}_{m-k}\mathbf{b} = \|\mathbf{Q}_{m-k}\mathbf{b}\|^2$$

$$\mathbf{J}_2(\mathbf{x}_2) = (\mathbf{1}_m - \mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k)\mathbf{J}_{n-k}(\mathbf{x}_2) - [\mathcal{D}(\mathbf{A})\mathbf{U}^{-1}\mathbf{Q}_k]^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2) \ ,$$
$$\mathbf{J}_2(\mathbf{x}_2) \approx (\mathbf{1}_m - \mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k)\mathbf{J}_{n-k}(\mathbf{x}_2) \ ,$$

$\mathbf{Q} \sim (m,m)$ is an orthogonal matrix, $\mathbf{U} \sim (k,k)$ is a regular upper-triangular matrix, or

$$\mathbf{A} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{U} \\ \mathbf{0} \end{pmatrix} \mathbf{P} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_k \\ \mathbf{0} \end{pmatrix} \mathbf{U}\mathbf{P} = \mathbf{Q}_k^{\mathrm{T}}\mathbf{U}\mathbf{P} \ ,$$

$$\mathbf{A}^+ = \mathbf{P}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{Q}_k \ ,$$
$$\mathbf{J}_2(\mathbf{x}_2) = (\mathbf{1}_m - \mathbf{Q}_k^{\mathrm{T}}\mathbf{Q}_k)\mathbf{J}_{n-k}(\mathbf{x}_2) - [\mathcal{D}(\mathbf{A})\mathbf{P}^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{Q}_k]^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2) \ ,$$

$\mathbf{P} \sim (k,k)$ is a permutation matrix¶.

If $r(\mathbf{A}) \equiv r < k$ (i.e. if the matrix $\mathbf{A}(\mathbf{x}_2)$ does not have full rank), then

$$\mathbf{A} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{U} & \mathbf{V} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{P} = \mathbf{Q}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_r \\ \mathbf{0} \end{pmatrix} (\, \mathbf{U} \quad \mathbf{V} \,)\, \mathbf{P} = \mathbf{Q}_r^{\mathrm{T}} (\, \mathbf{U} \quad \mathbf{V} \,)\, \mathbf{P} \ ,$$

$$\mathbf{A}^+ = \mathbf{P}^{\mathrm{T}} \begin{pmatrix} \mathbf{U}^{-1} \\ \mathbf{0} \end{pmatrix} \mathbf{Q}_r = \mathbf{P}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_r \\ \mathbf{0} \end{pmatrix} \mathbf{U}^{-1}\mathbf{Q}_r = \mathbf{P}_r^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{Q}_r \ ,$$
$$\mathbf{f}_2(\mathbf{x}_2) = (\mathbf{1}_m - \mathbf{Q}_r^{\mathrm{T}}\mathbf{Q}_r)\mathbf{b} \ ,$$
$$\|\mathbf{f}_2(\mathbf{x}_2)\|^2 = \|\mathbf{Q}_{m-r}\mathbf{b}\|^2 \ ,$$
$$\mathbf{J}_2(\mathbf{x}_2) = (\mathbf{1}_m - \mathbf{Q}_r^{\mathrm{T}}\mathbf{Q}_r)\mathbf{J}_{n-k}(\mathbf{x}_2) - [\mathcal{D}(\mathbf{A})\mathbf{P}_r^{\mathrm{T}}\mathbf{U}^{-1}\mathbf{Q}_r]^{\mathrm{T}}\mathbf{f}_2(\mathbf{x}_2) \ ,$$
$$\mathbf{J}_2(\mathbf{x}_2) \approx (\mathbf{1}_m - \mathbf{Q}_r^{\mathrm{T}}\mathbf{Q}_r)\mathbf{J}_{n-k}(\mathbf{x}_2) \ ,$$

$\mathbf{Q} \sim (m,m)$ is an orthogonal matrix, $\mathbf{U} \sim (r,r)$ is a regular upper-triangular matrix, $\mathbf{P} \sim (k,k)$ is a permutation matrix¶.

1. Golub G. H., Pereyra V., *SIAM J. Numer. Anal.* **10** (1973) 413
2. Maeder M., Zuberbühler A. D., *Anal. Chem.* **62** (1990) 2220

¶ Formally the same relations we get if we substitute the singular value decomposition for the QR-decomposition with column pivoting. If $r(\mathbf{A}) = k$, then

$$\mathbf{A} = \mathbf{U}^{\mathrm{T}} \begin{pmatrix} \mathbf{\Sigma} \\ \mathbf{0} \end{pmatrix} \mathbf{V} = \mathbf{U}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_k \\ \mathbf{0} \end{pmatrix} \mathbf{\Sigma}\mathbf{V} = \mathbf{U}_k^{\mathrm{T}}\mathbf{\Sigma}\mathbf{V} \ ,$$

$$\mathbf{A}^+ = \mathbf{V}^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{U}_k \ ,$$

$\mathbf{U} \sim (m,m)$ a $\mathbf{V} \sim (k,k)$ are orthogonal matrices, $\mathbf{\Sigma} \equiv \mathrm{diag}\,(\sigma_1, \ldots, \sigma_k) \sim (k,k)$ is a diagonal matrix with positive diagonal elements $\sigma_1 > \ldots > \sigma_k > 0$. If $r(\mathbf{A}) \equiv r < k$, then

$$\mathbf{A} = \mathbf{U}^{\mathrm{T}} \begin{pmatrix} \mathbf{\Sigma} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{V} = \mathbf{U}^{\mathrm{T}} \begin{pmatrix} \mathbf{1}_r \\ \mathbf{0} \end{pmatrix} (\, \mathbf{\Sigma} \quad \mathbf{0} \,)\, \mathbf{V} = \mathbf{U}_r^{\mathrm{T}}\mathbf{\Sigma}\, (\, \mathbf{1}_r \quad \mathbf{0} \,)\, \mathbf{V} = \mathbf{U}_r^{\mathrm{T}}\mathbf{\Sigma}\mathbf{V}_r \ ,$$

$$\mathbf{A}^+ = \mathbf{V}_r^{\mathrm{T}}\mathbf{\Sigma}^{-1}\mathbf{U}_r \ ,$$

$\mathbf{U} \sim (m,m)$ a $\mathbf{V} \sim (k,k)$ are orthogonal matrices, $\mathbf{\Sigma} \equiv \mathrm{diag}\,(\sigma_1, \ldots, \sigma_r) \sim (r,r)$ is a diagonal matrix with positive diagonal elements $\sigma_1 > \ldots > \sigma_r > 0$.