A. Quarteroni, R. Sacco, F. Saleri,

Numerical Mathematics

2nd Edition. TAM37. Springer, Berlin 2007

Errata

April 16, 2013

pag. 30: Hint for the solution of Exercise n. 4: Note that $\mathrm{I} + \mathrm{B} = 2\mathrm{I} - (\mathrm{I} - \mathrm{B})$

pag. 34, row 5: "correspondance" has to be replaced by "correspondence"

pag. 75, row -5: "$i \neq j$" has to be replaced by "$i \leq j$"

pag. 78, row 1: "*If A is a matrix diagonally dominant...*" has to be replaced by "*If A is a matrix strictly diagonally dominant...*"

pag. 86, row 5: "(that is, $\tilde{\mathrm{Q}}^{-1} = \tilde{\mathrm{Q}}^T$)" has to be replaced by "(that is, $\tilde{\mathrm{Q}}^T\tilde{\mathrm{Q}} = \mathrm{I}_n$, being I_n the identity matrix of size n)"

pag. 140: After formula (4.29) add:
"where in this case $K_2(\mathrm{P}^{-1}\mathrm{A}) = \lambda_1/\lambda_n$."

pag. 141, row 1: "The matrix R_α is symmetric positive definite" has to be replaced by "The matrix R_α is symmetric"

pag. 155: After formula (4.49) add:
"where $K_2(\mathrm{A}) = \lambda_1/\lambda_n$ and λ_1 (λ_n, resp.) is the maximum (minimum, resp.) eigenvalue of A."

pag. 255, row -9,-8: the phrase "Assuming that $f \in C^1(\mathcal{I})$ and that $f'(\alpha) \neq 0$ (i.e., α is a simple root of f),"
has to be replaced by
"Assuming that $f \in C^1(\mathcal{J})$ and that $f'(x) \neq 0,\ \forall x \in \mathcal{J} \setminus \{\alpha\}$"

pag. 255, row -2,-1: the phrase "... by a higher order of convergence, Newton's method being of order 2"
has to be replaced by
"... by a higher order of convergence when α is a simple root of f (i.e. $f'(\alpha) \neq 0$). As a matter of fact, in this case the Newton's method is of order 2"

pag. 262, row 7,8: the phrase "... ϕ, which is continuous and differentiable in a neighborhood $\mathcal{J}$ of α."
has to be replaced by
"...ϕ, which is continuous and continuously differentiable in a neighborhood $\mathcal{J}$ of α."

pag. 262, row 11,12: "$x^{(n)}$" has to be replaced by "$x^{(k)}$"
"$x^{(n+1)}$" has to be replaced by "$x^{(k+1)}$"

pag. 264, row 4: "then the method (6.16) is no longer second-order convergent." has to be replaced by
"(i.e., $f'(\alpha) = 0, \ \ldots, f^{(m-1)}(\alpha) = 0$), then the Newton method (6.16) still converges, under the condition that $x^{(0)}$ is properly chosen and $f'(x) \neq 0 \ \forall x \in \mathcal{J} \setminus \{\alpha\}$, but now it is only first-order convergent."

pag. 283, exercise 6: " Analyze the convergence of the fixed-point method $x^{(k+1)} = \phi_j(x^{(k)})$ for computing the zeros $\alpha_1 = -1$ and $\alpha_2 = 2$ of the function $f(x) = x^2 - x - 2$, when the following iteration functions are used: $\phi_1(x) = x^2 - 2$, $\phi_2(x) = \sqrt{2+x}$ $\phi_3(x) = -\sqrt{2+x}$ and $\phi_4(x) = 1 + 2/x$, $x \neq 0$.
[*Solution*: the method is non convergent with ϕ_1, it converges only to α_2, with ϕ_2 and ϕ_4, while it converges only to α_1 with ϕ_3]."
has to be replaced by :
"Analyze the behaviour (about either consistency and convergence) of the fixed-point method $x^{(k+1)} = \phi_j(x^{(k)})$ for computing the zeros $\alpha_1 = -1$ and $\alpha_2 = 2$ of the function $f(x) = x^2 - x - 2$, when the following iteration functions are used: $\phi_1(x) = x^2 - 2$, $\phi_2(x) = \sqrt{2+x}$, $\phi_3(x) = -\sqrt{2+x}$ e $\phi_4(x) = 1 + 2/x$, $x \neq 0$.

[*Solution*: the method with ϕ_1 and ϕ_4 is consistent in order to compute both roots of f, but with ϕ_2 it is consistent only in correspondence of α_2, while with ϕ_3 only in correspondence of α_1. The choice ϕ_1 is not convergent, the choices ϕ_2 and ϕ_3 are convergent, while ϕ_4 is convergent only to α_2.]"

pag. 304, row -19: "`fmins`" has to be replaced by "`fminsearch`"

pag. 304, row -10: "`fmins`" has to be replaced by "`fminsearch`"

pag. 321, row 6: "`fmins`" has to be replaced by "`fminsearch`"

pag. 324, row 13: "`fmins`" has to be replaced by "`fminsearch`"

pag. 393, row 15: "(9.26) immediately follows." has to be replaced by "(9.26) immediately follows, recalling that $\gamma_n = n + 2$."

pag. 401, row -12: "$\displaystyle\int_0^\pi (e^{x/2} + \cos 4x)dx = 2(e^\pi - 1) \simeq 7.621,$"
has to be replaced by
"$\displaystyle\int_0^\pi (e^{x/2} + \cos 4x)dx = 2(e^{\pi/2} - 1) \simeq 7.621,$"

pag. 447, row -1:
"$$\Pi_N^F f(x_j) = \sum_{k=0}^{N-1} \widetilde{f}_k e^{ikjh} e^{-ijh\frac{N}{2}} = \sum_{l=0}^{N-1} f(x_l) \left[\frac{1}{N} \sum_{k=0}^{N-1} e^{-ik(l-j)h}\right] = f(x_j).$$"
has to be replaced by
"$$\Pi_N^F f(x_j) = \sum_{k=0}^{N-1} \widetilde{f}_k e^{ikjh} e^{-ijh\frac{N}{2}} = \sum_{l=0}^{N-1} f(x_l) \left[\frac{1}{N} \sum_{k=0}^{N-1} e^{-ik(l-j)h} e^{i\pi(l-j)}\right] =$$
$f(x_j)$."

pag. 448, formula (10.55):
"$$f(x_j) = \sum_{k=0}^{N-1} \widetilde{f}_k e^{ik(j-\frac{N}{2})h} = \sum_{k=0}^{N-1} \widetilde{f}_k W_N^{-(j-\frac{N}{2})k}, \; j = 0, \dots, N-1.$$"
has to be replaced by
"$$f(x_j) = \sum_{k=0}^{N-1} \widetilde{f}_k e^{ij(k-\frac{N}{2})h} = \sum_{k=0}^{N-1} \widetilde{f}_k W_N^{-(k-\frac{N}{2})j}, \; j = 0, \dots, N-1.$$"

pag. 448, row 9:
"$C_{jk} = W_N^{-(j-\frac{N}{2})k}, \quad j,k = 0,\dots,N-1.$"
has to be replaced by
"$C_{jk} = W_N^{-(k-\frac{N}{2})j}, \quad j,k = 0,\dots,N-1.$"

pag. 449: Program 89 correctly works, nevertheless the variables `j` and `k` could be exchanged in order they assume the meaning assigned in the definition of matrix C at pag. 448. The new version reads:

```
function fv = idft(N,fc)
%IDFT Inverse discrete Fourier transform.
%   FV=IDFT(N,F) computes the coefficients of the
%   inverse discrete Fourier transform of a function F.
h  = 2*pi/N; wn = exp(-i*h);
for j=0:N-1
   s = 0;
   for k=0:N-1
      s = s + fc(k+1)*wn^(-j*(k-N/2));
   end
   fv (j+1) = s;
end
return
```

pag. 477, exercise 5: "Compute weights and nodes of the following quadrature formulae

$$\int_a^b w(x)f(x)dx = \sum_{i=0}^{n} \omega_i f(x_i),$$

in such a way that the order is maximum, setting

$$\begin{array}{ll} \omega(x) = \sqrt{x}, & a = 0, \quad b = 1, n = 1; \\ \omega(x) = 2x^2 + 1, & a = -1, b = 1, n = 0; \\ \omega(x) = \begin{cases} 2 \text{ if } & 0 < x \leq 1, \\ 1 \text{ if } & -1 \leq x \leq 0 \end{cases} & a = -1, b = 1, n = 1. \end{array}$$

[*Solution*: for $\omega(x) = \sqrt{x}$, the nodes $x_1 = \frac{5}{9} + \frac{2}{9}\sqrt{10/7}$, $x_2 = \frac{5}{9} - \frac{2}{9}\sqrt{10/7}$ are obtained, from which the weights can be computed (order 3); for $\omega(x) = 2x^2 + 1$, we get $x_1 = 3/5$ and $\omega_1 = 5/3$ (order 1); for $\omega(x) =$

$2x^2+1$, we have $x_1 = \frac{1}{22} + \frac{1}{22}\sqrt{155}$, $x_2 = \frac{1}{22} - \frac{1}{22}\sqrt{155}$ (order 3).]"
has to be replaced by
"Compute weights α_i and nodes x_i of the following quadrature formulae

$$\int_a^b w(x)f(x)dx = \sum_{i=0}^{n} \alpha_i f(x_i),$$

in such a way that the order is maximum, setting

$$\text{(A) } w(x) = \sqrt{x}, \qquad a=0, \quad b=1,\ n=1;$$

$$\text{(B) } w(x) = 2x^2+1, \qquad a=0, \quad b=1,\ n=0;$$

$$\text{(C) } w(x) = \begin{cases} 2 \ \text{ if } \ 0 < x \le 1, \\ 1 \ \text{ if } \ -1 \le x \le 0 \end{cases} \qquad a=-1,\ b=1,\ n=1.$$

[*Solution*: case (A): the nodes $x_0 = \frac{5}{9} + \frac{2}{9}\sqrt{10/7}$, $x_1 = \frac{5}{9} - \frac{2}{9}\sqrt{10/7}$ are obtained, from which the weights α_i can be computed (order 3); case (B) we get $x_0 = 3/5$ and $\alpha_0 = 5/3$ (order 1); case (C) we get $x_0 = \frac{1}{22} + \frac{1}{22}\sqrt{155}$, $x_1 = \frac{1}{22} - \frac{1}{22}\sqrt{155}$ (order 3).]"

pag. 484,485: The definition 11.4 (Zero-stability of one-step methods) has to be replaced by
Definition 11.4 (Zero-stability of one-step methods) The numerical method (11.11) for the approximation of problem (11.1) is *zero-stable* if
$\exists h_0 > 0$, $\exists C > 0$ and $\exists \varepsilon_0 > 0$ such that $\forall h \in (0, h_0]$ and $\forall \varepsilon \in (0, \varepsilon_0]$, if $|\delta_n| \le \varepsilon$, $0 \le n \le N_h$, then

$$|z_n^{(h)} - u_n^{(h)}| \le C\varepsilon, \qquad 0 \le n \le N_h,$$

where $z_n^{(h)}$ and $u_n^{(h)}$ are respectively the solutions of the problems

$$\begin{cases} z_{n+1}^{(h)} = z_n^{(h)} + h\left[\Phi(t_n, z_n^{(h)}, f(t_n, z_n^{(h)}); h) + \delta_{n+1}\right],\ n = 0, \ldots, N_h - 1 \\ z_0^{(h)} = y_0 + \delta_0, \end{cases}$$

$$\begin{cases} u_{n+1}^{(h)} = u_n^{(h)} + h\Phi(t_n, u_n^{(h)}, f(t_n, u_n^{(h)}); h),\ n = 0, \ldots, N_h - 1 \\ u_0^{(h)} = y_0. \end{cases}$$

pag. 489, row -9: "fot" has to be replaced by "for"

pag. 498, row -2: The phrase "Notice that the LTE is exactly $\mathcal{L}[y(t_n); h]$." has to be replaced by "Notice that the LTE is exactly $\frac{1}{h}\mathcal{L}[y(t_n); h]$."

pag. 499, row 2,3: The rows "Consequently, if the MS method has order q and $y \in C^{q+1}(I)$, we obtain

$$\tau_{n+1}(h) = C_{q+1} h^{q+1} y^{(q+1)}(t_{n-p}) + \mathcal{O}(h^{q+2}).''$$

have to be replaced by
"Consequently, if

$$C_0 = C_1 = \ldots = C_q = 0, \tag{0.1}$$

then

$$\mathcal{L}[y(t_n); h] = h\tau_{n+1}(h) = C_{q+1} h^{q+1} y^{(q+1)}(t_{n-p}) + \mathcal{O}(h^{q+2}).$$

In view of Definition 11.9, the MS method is of order q. Note that conditions (0.1) give rise to algebraic conditions on the MS coefficients $\{a_j, b_j\}$, as we will see in Theorem 11.3 (pag. 503). It is worth noticing that a different choice of the origin about which the terms $w(t - jh)$ and $w'(t - jh)$ are expanded would yield an a-priori different set of constants $\{C_k\}$. However, as pointed out in [Lam91, pp.48–49] the first non-vanishing coefficient C_{q+1} is invariant (whereas the other coefficients C_{q+j}, $j \geq 2$, are not)."

pag. 505: The first three rows of Definition 11.13 have to be replaced by:
The multistep method (11.45) is zero-stable if
$\exists h_0 > 0$, $\exists C > 0$ and $\exists \varepsilon_0 > 0$ such that $\forall h \in (0, h_0]$, $\forall \varepsilon \in (0, \varepsilon_0]$, if $|\delta_n| \leq \varepsilon$, $0 \leq n \leq N_h$, then

pag. 509, row 8,9:

$$u_n = \sum_{j=1}^{k'} \left(\sum_{s=0}^{m_j - 1} \gamma_{sj} n^s \right) [r_j(h\lambda)]^n, \qquad n = 0, 1, \ldots,$$

where $r_j(h\lambda)$, $j = 1, \ldots, k'$,
have to be replaced by

$$u_n = \sum_{j=0}^{k'} \left(\sum_{s=0}^{m_j-1} \gamma_{sj} n^s \right) [r_j(h\lambda)]^n, \qquad n = 0, 1, \dots,$$

where $r_j(h\lambda)$, $j = 0, \dots, k'$,

pag. 511, formula (11.66): Formula (11.66) has to be replaced by

$$\exists h_0 > 0, \exists C > 0 \,:\, \forall h \le h_0 \quad |u_n| \le C(|u_0| + \ldots + |u_p|), \quad \forall n \ge p+1.$$

pag. 521, row 3: The phrase "u_{n+1}, assuming" has to be replaced by "u^*_{n+1}, obtained assuming"

pag. 521, row 9: The formula

$$u_{n+1} = y_n + hF(t_n, y_n, h; f) = y_n + h(b_1 K_1 + b_2 K_2)$$

has to be replaced by

$$u^*_{n+1} = y_n + hF(t_n, y_n, h; f) = y_n + h(b_1 K_1 + b_2 K_2)$$

pag. 521, row 17: The formula

$$u_{n+1} = y_n + hf_n(b_1 + b_2) + h^2 c_2 b_2 (f_{n,t} + f_n f_{n,y}) + \mathcal{O}(h^3)$$

has to be replaced by

$$u^*_{n+1} = y_n + hf_n(b_1 + b_2) + h^2 c_2 b_2 (f_{n,t} + f_n f_{n,y}) + \mathcal{O}(h^3)$$

pag. 522: Starting from formula (11.74) and up to formula (11.76) (included),
"u_{n+1}" has to be replaced by "u^*_{n+1}",
"$\widehat{u}_{n+1}$" has to be replaced by "$\widehat{u}^*_{n+1}$".

pag. 522: After formula

$$y_{n+1} - u^*_{n+1} \simeq \frac{u^*_{n+1} - \widehat{u}^*_{n+1}}{(2^{p+1} - 1)} = \mathcal{E}$$

add the following phrase:
In practice, since both u^*_{n+1} and $\widehat{u}^*_{n+1}$ are unknowns, we evaluate

$$\mathcal{E} = \frac{u_{n+1} - \widehat{u}_{n+1}}{(2^{p+1} - 1)},$$

u_{n+1} and $\widehat{u}_{n+1}$ being the numerical solutions obtained with stepsize h and $2h$, rispectively.

pag. 585, row -1: The word "endspores" has to be replaced by "endospores"

pag. 593, row 1: the phrase "(see Exercise 3)" has to be replaced by : "(see Exercise 3 of Chapter 4. Note that $h^2\mathrm{A}_{\mathrm{fd}}$ coincides with the matrix A of Exercise 3 of Chapter 4 with $\alpha = 2$)"

pag. 595, row 11-15: The phrase "Thus, its Cholesky decomposition $\mathrm{K} = \mathrm{H}^T\mathrm{H}$ where H is upper triangular (see Section 3.4.2) can be carried out at $t = 0$. Consequently, at each time step the following two linear triangular systems, each of size equal to N_h, must be solved, with a computational cost of $N_h^2/2$ flops"
has to be replaced by
"Moreover, K is independent of k and then it can be factorized once at $t = 0$. For the one-dimensional case that we are handling, this factorization is based on the Thomas method (see Section 3.7.2) and it requires a number of operation proportional to N_h. In the multidimensional case the use of the Cholesky factorization $\mathrm{K} = \mathrm{H}^T\mathrm{H}$, H being an upper triangular matrix (see Section 3.4.2), will be more convenient. Consequently, at each time step the following two linear triangular systems, each of size equal to N_h, must be solved:"

pag. 611, row 4,5: In Table 13.1:
"$\mathcal{O}(\Delta x^2/\Delta t+\Delta t+\Delta x)$" has to be replaced by "$\mathcal{O}(\Delta x^2/\Delta t+\Delta t+\Delta x^2)$"
"$\mathcal{O}(\Delta t^2 + \Delta x^2)$" has to be replaced by "$\mathcal{O}(\Delta t^2 + \Delta x^2 + \Delta t\Delta x^2)$"

pag. 611, row -13: the phrase "and so is the artificial viscosity" has to be dropped.

pag. 612, row 6:

$$\tau_j^n = \frac{u(x_j, t^{n+1}) - u(x_j, t^n)}{\Delta t} - a\frac{u(x_{j+1}, t^n) - u(x_{j-1}, t^n)}{2\Delta x}$$

has to be replaced by

$$\tau_j^n = \frac{u(x_j, t^{n+1}) - u(x_j, t^n)}{\Delta t} + a\frac{u(x_{j+1}, t^n) - u(x_{j-1}, t^n)}{2\Delta x}$$

pag. 612, row 12: “for suitable integers p and q” has to be replaced by “for suitable positive p and q”

pag. 619, row 13,14: formulas ”$l = 10\Delta x$” and ”$l = 4\Delta x$” have to be replaced by ”$l = 20\Delta x$” and ”$l = 8\Delta x$”, respectively.

pag. 632, row 1: “Computed solutions using the NM...” has to be replaced by “Computed solutions using the NW...”

pag. 638: “Erdös P. (1961) Problems and Results on the Theory of Interpolation. *Acta Math. Acad. Sci. Hungar.* 44: 235–244.”
has to be replaced by
“Erdös P. (1961) Problems and Results on the Theory of Interpolation. II. *Acta Math. Acad. Sci. Hungar.* 12: 235–244.”