Pade approximation of Gauss function

Tomasz Chwiej

28th January 2025

1 Itroduction

Known smooth function $f(x)$ can be approximated with rational function $R_{N,M}$

$$R_{N,M}(x) = \frac{P_N(x)}{Q_M(x)} = \frac{\sum\limits_{i=0}^{N} a_i x^i}{\sum\limits_{i=0}^{M} b_i x^i}, \qquad (b_0 = 1) \tag{1}$$

In order to find the coefficients of polynomials P_N and Q_N we replace $f(x)$ with Maclaurin series (for $x = 0$)

$$f(x) = \sum_{k=0}^{\infty} c_k x^k \tag{2}$$

and compare the derivatives of $f(x)$ and $R_{N,M}(x)$ sequentially for $k = 0, 1, \ldots, N + M$ order

$$\left.\frac{d^k R_{N,M}(x)}{dx^k}\right|_{x=0} = \left.\frac{d^k f(x)}{dx^k}\right|_{x=0} \tag{3}$$

These conditions generate system of linear equations (SLE)

$$\sum_{m=1}^{N} b_m \cdot c_{N-m+k} = -c_{N+k}, \quad k = 1, 2, \ldots, N \tag{4}$$

which has the following matrix form

$$\begin{bmatrix} c_{N-M+1} & c_{N-M+2} & \cdots & c_N \\ c_{N-M+2} & c_{N-M+3} & \cdots & c_{N+1} \\ \vdots & \vdots & \ddots & \vdots \\ c_N & c_{N+1} & \cdots & c_{N+M-1} \end{bmatrix} \begin{bmatrix} b_M \\ b_{M-1} \\ \vdots \\ b_1 \end{bmatrix} = \begin{bmatrix} -c_{N+1} \\ -c_{N+2} \\ \vdots \\ -c_{N+M} \end{bmatrix} \tag{5}$$

and must be solved so as to find coefficients $\vec{b} = [b_0, b_1, \ldots, b_M]$.
Next we use the relation

$$a_i = \sum_{j=0}^{i} c_{i-j} \cdot b_j, \quad i = 0, 1, \ldots, N \tag{6}$$

in order to determine the coefficents $\vec{a} = [a_0, a_1, \ldots, a_N]$.

2 Practical part

1. Write the computer program which will compute Pade approximation of Gauss function

$$f(x) = \exp(-x^2) \tag{7}$$

for the following P_N and Q_M polynomials' degrees $(N, M) = (2, 2),\ (4, 4),\ (6, 6),\ (2, 4),\ (2, 6),\ (2, 8)$.

   - Coefficients of Maclaurin series (c_k) calculate directly from series expansion of exponential function $f(x) = \exp(-x^2)$

$$\exp(-x^2) = \sum_{p=0}^{\infty} (-1)^p \frac{x^{2p}}{p!} = \sum_{k=0}^{\infty} c_k \cdot x^k \tag{8}$$

     Values of coefficents c_k save in vector $\vec{c} = [c_0, c_1, \ldots, c_n]$, $\quad n = N + M$.

   - Solve system of linear equations with routine **LAPACKE_dsysv()** (see Sec.3). SLE is defined in (Eq. 5)

$$A\vec{x} = \vec{y} \tag{9}$$

     where:

$$A_{i,j} = c_{N-M+i+j+1}, \quad i, j = 0, 1, \ldots, M-1 \tag{10}$$

$$y_i = -c_{N+1+i}, \quad i = 0, 1, \ldots, M-1 \tag{11}$$

     Solution of SLE given in Eq. 9 contains coefficents of polynomial $Q_M(x)$

$$b_0 = 1 \quad \text{and} \quad b_{M-i} = x_i, \quad i = 0, 1, \ldots, M-1 \tag{12}$$

     Save these coefficents in vector $\vec{b} = [b_0, b_1, \ldots, b_M]$.

   - Next, determine the coefficents of polynomial $P_N(x)$ accordingly with Eq. 6. Save the coefficients in vector $\vec{a} = [a_0, a_1, \ldots, a_N]$.

2. For each pair of indices (N, M) prepare separate figure showing the functons $f(x)$ and $R_{N,M}(x)$ (Eq. 1) over an interval $x \in [-5, 5]$.

3. At home prepare report. Comment on the quality of approximation and try to answear to questions:

   - Why the rational function $R_{N,M}$ can not approximate well the Gauss function for $N = M$?
   - To what value tends asymptotically $R_{6,6}$ for $x \to \pm\infty$?

3 Hints

1. In Eq. 8 occurs factorial function, to compute its value you may use the follownig recursive function

```
double function  factorial(double n)
        if n = 0 then
                return 1
        else
                return n*factorial(n-1)
        end if
end function
```

2. The matrix A defined in Eq. 9 is symmetric, therefore to solve SLE with this matrix use the routine **LAPACKE_dsysv()**

```
lapack_int LAPACKE_dsysv (int matrix_layout, char uplo, lapack_int n,
        lapack_int nrhs, double *A, lapack_int lda, lapack_int *ipiv,
        double *b, lapack_intldb);
```

with following arguments: **matrix_layout=LAPACK_COL_MAJOR**, **uplo='U'** - procedure assume the upper triangle part of A is filled, **n=M** - number of equations, **nrhs=1** - number of right hand sides, **A** - 1D array of size M^2 contains the elements of matrix A filled in column-major order, **lda=M** - leading dimension of matrix A, **ipiv** - 1D array of integers of size M, **b** - 1D array of size M (right hand side of SLE), **ldb=M** - leading dimension of array **b**.

Since we assume column-major order of elements in matrix A (represented by 1D array), this shall be filled as follows

```
for i=0 to M-1 by 1 do
        for j=0 to M-1 by 1 do
                A[j + i * M] ← c[N − M + i + j + 1]
        end do
end do
```

4 Example results

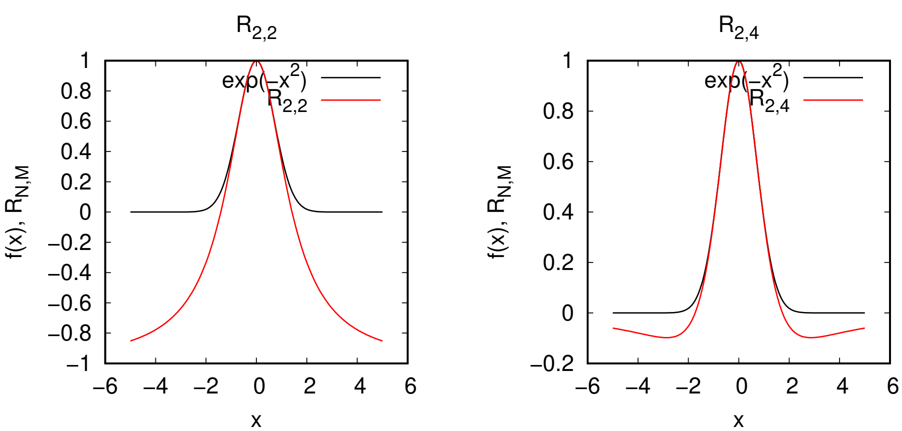

Figure 1: Pade approximation $R_{N,M}(x)$ of Gauss function $f(x) = e^{-x^2}$ for $R_{2,2}$ (left) and $R_{2,4}$ (right).