Problem 0.0.1

The exponential integrals are the functions E_n defined by

$$E_n(x) = \int_1^\infty (e^{xt}t^n)^{-1}\,\mathrm{d}t, \quad (n \geqslant 0, x > 0).$$

These functions satisfy the equation

$$nE_{n+1}(x) = e^{-x} - xE_n(x).$$

If $E_1(x)$ is known, can this computation be used to compute $E_2(x), E_3(x), \ldots$ accurately?

Solution

In general, no. We first rewrite the equation as

$$E_{n+1}(x) = \frac{e^{-x}}{n} - \frac{x}{n}E_n(x).$$

This is relatively accurate for small x. When x is small, $x/n < 1$ for almost n's and the error of δ units in $E_n(x)$ gradually decays to 0.

For larger x's, however, every time when we compute $E_{n+1}(x)$ for $n < x$, the error is multiplied by a factor > 1 as $x/n > 1$ when $n < x$. For example, when $x = 10$, the initial error δ caused by rounding $E_1(x)$ becomes a stunning

$$\frac{10}{1} \cdot \frac{10}{2} \cdot \cdots \cdot \frac{10}{9} \cdot \delta = \frac{10^9\delta}{9!} \approx 2755\delta,$$

which greatly distrusts the accuracy. In fact, when trying $x = 15$ on my MATLAB with a machine epsilon 2^{-52}, $E_{12}(x)$ returns a negative value, which is absurd.

Problem 0.0.2

Show that the following method has third-order convergence for computing $\sqrt{R}$:

$$x_{n+1} = \frac{x_n(x_n^2 + 3R)}{3x_n^2 + R}.$$

Proof. Let $e_n = x_n - \sqrt{R}$ be the error. Then,

$$\begin{aligned} e_{n+1} = x_{n+1} - \sqrt{R} &= \frac{x_n(x_n^2+3R)}{3x_n^2+R} - \sqrt{R} \\ &= \frac{x_n^3 + 3Rx_n - 3x_n^2\sqrt{R} - R\sqrt{R}}{3x_n^2+R} \\ &= \frac{(x_n - \sqrt{R})^3}{3x_n^2+R} = \frac{e_n^3}{3x_n^2+R}. \end{aligned}$$

The denominator shows that this method indeed has third-order convergence. □

Problem 0.0.3: (a)

(Problem 4.7.1, p.217) Prove that if A is symmetric, then the gradient of the function

$$q(x) = \langle x, Ax\rangle - 2\langle x, b\rangle$$

at x is $2(Ax - b)$.

Proof. First notice that the partial derivative

$$g := \frac{\partial \langle x, y\rangle}{\partial x} = y. \tag{1}$$

Indeed, consider g_1 (the first component of g). It is defined as

$$\frac{\partial}{\partial x_1}\sum_{i=1}^{n} x_i y_i = y_1,$$

and likewise for all the remaining ones. Vector-wise we also have the chain rule:

$$\begin{aligned}\frac{\mathrm{d}\langle x, Ax\rangle}{\mathrm{d}x} &= \frac{\mathrm{d}x^T}{\mathrm{d}x}\cdot\frac{\partial\langle x, Ax\rangle}{\partial x} + \frac{\mathrm{d}(Ax)^T}{\mathrm{d}x}\cdot\frac{\partial\langle x, Ax\rangle}{\partial(Ax)}\\ &= Ax + \frac{\mathrm{d}(x^T A^T)}{\mathrm{d}x}\cdot x\\ &= Ax + A^T x = (A + A^T)x = 2Ax.\end{aligned}$$

From (1), it is also clear that $\partial(2\langle x, b\rangle)/\partial x = 2b$. Therefore the gradient of $q(x) = 2(Ax - b)$. □

Problem 0.0.3: (b)

(Problem 4.7.6, p.218) Prove that if $\hat{t}$ is defined by

$$\hat{t} = \frac{\langle v, b - Ax\rangle}{\langle v, Av\rangle}$$

and if $y = x + \hat{t}v$, then $v \perp (b - Ay)$, that is, $\langle v, b - Ay\rangle = 0$.

Proof. Indeed, this result follows directly from brute-force computation:

$$\begin{aligned}\langle v, b - Ay\rangle &= \langle v, b - Ax\rangle - \left\langle v, A\hat{t}v\right\rangle\\ &= \frac{\langle v, b - Ax\rangle}{\langle v, Av\rangle}\cdot\langle v, Av\rangle - \left\langle v, A\hat{t}v\right\rangle\\ &= \hat{t}\langle v, Av\rangle - \left\langle v, A\hat{t}v\right\rangle = \left\langle v, A\hat{t}v\right\rangle - \left\langle v, A\hat{t}v\right\rangle = 0.\end{aligned}$$

□