problem sheet 12

discussion: week of Monday, 17.1.2022

12.1. The *secant method* (i.e., Broyden's method in 1D) to find the zero x^* of $F(x) = 0$ is defined as follows given initial points x_0, x_1:

$$x_{n+1} = x_n - \frac{x_n - x_{n-1}}{F(x_n) - F(x_{n-1})} F(x_n), \qquad n = 1, 2, \ldots$$

(If $F(x_n) = F(x_{n-1})$ the difference quotient is formally replaced with $F'(x_n)$.) Let $F(x) = 2 - x^2 - e^x$.

1. Compute, using Newton's method the positive zero x^* of F to machine precision.
2. Compute the zero x^* with the secant method. Set $x_1 = x_0 - \frac{F(x_0)}{F'(x_0)}$. Plot for $n \in \{1, \ldots, 8\}$ and $x_0 = 2.5$ the error $|x^* - x_n|$ versus the step number n. Also plot the *numerical convergence order* $p_n = \log(|x^* - x_{n+1}|)/\log(|x^* - x_n|)$ versus n. What convergence order do you observe?
3. Compare the *efficiency* of the secant method with that of the Newton method by comparing accuracy versus number of function evaluations. To that end, assume that a Newton step costs 3 function evaluations (this is realistic assuming that F' is approximated with a difference quotient) and plot achieved accuracy versus number of function evaluations. Which method is more efficient?

12.2. Show the following convergence result for the inverse iteration with shift: Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be diagonalizable and $\lambda \in \mathbb{R}$. Let the eigenvalues of $\mathbf{A}$ be numbered such that $|\lambda_1 - \lambda| \geq |\lambda_2 - \lambda| \geq \cdots |\lambda_{n-1} - \lambda| > |\lambda_n - \lambda|$. Then there exists $C > 0$ such that there holds for the approximations $\widetilde{\lambda}_\ell$ of the inverse iteration:

$$|\lambda_n - \widetilde{\lambda}_\ell| \leq C \left| \frac{\lambda_n - \lambda}{\lambda_{n-1} - \lambda} \right|^\ell, \qquad \ell = 0, 1, \ldots,$$

12.3. Consider the vector iteration ("power method") for the matrix $\mathbf{A}$ and the following three initial vectors $x_0^{(j)}$, $j = 0, 1, 2$:

$$\mathbf{A} = \begin{pmatrix} 2 & \\ & -2 \end{pmatrix}, \qquad \mathbf{x}_0^{(1)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{x}_0^{(2)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mathbf{x}_0^{(3)} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Discuss the behavior of the vector iteration. Do the eigenvalue approximations $\widetilde{\lambda}_\ell$ and the iterates $\mathbf{x}_\ell$ converge? If so, what do they converge to?

12.4. Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a *symmetric* matrix and define the Rayleigh quotient

$$R(\mathbf{x}) := \frac{\mathbf{x}^\top \mathbf{A}\mathbf{x}}{\mathbf{x}^\top \mathbf{x}}$$

a) Show that the largest eigenvalue λ_1 is the maximum of R and that smallest eigenvalue λ_n the minimum of R.

b) The minimization property of λ_n suggests to use descent methods. Formulate such a method. To that end, check that

$$\mathbf{g} := \nabla R(\mathbf{x}) = 2\frac{\mathbf{A}\mathbf{x}\|\mathbf{x}\|_2^2 - (\mathbf{x}^\top \mathbf{A}\mathbf{x})\mathbf{x}}{\|\mathbf{x}\|^4}.$$

Use the *Armijo rule* to determine the step length.
Remark: in the present case, it is possible to solve the 1-dimensional minimization problem, i.e., finding t with $R(\mathbf{x} + t\mathbf{g}) = \min_\tau R(\mathbf{x} + \tau\mathbf{g})$, explicitly.

c) Use your algorithm to find the smallest eigenvalue of the tridiagonal matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ given by

$$\mathbf{A} = \frac{1}{h^2} \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & & \ddots & -1 \\ & & & -1 & 2 \end{pmatrix}, \qquad h = \frac{1}{n}.$$

Plot (`semilogy`) the error versus the iteration number ℓ, where $\ell = 1, \ldots, 100$. You may compute the exact value of the eigenvalue using `matlab`'s `eig` command. What do you observe in dependence on $n \in \{10, 20, 40, 80\}$?

Remark: As in Problem 11.1 the matrix $\mathbf{A}$ is obtained as a discretization of $-\frac{d^2u}{dx^2}$.