MA2501, Spring 2020, Numerical Methods

February 6, 2020

Problem 1

a) The vector function $\mathbf{x} \mapsto \mathbf{f}(\mathbf{x})$ of two variables is defined by

$$f_1(x_1, x_2) = x_1^2 + x_2^2 - 2, \quad f_2(x_1, x_2) = x_1 - x_2.$$

Verify that the equation $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ has two solutions $x_1 = x_2 = 1$ and $x_1 = x_2 = -1$. Show that one iteration of Newton's method for the solution of this system gives $\mathbf{x}^{(1)} = (x_1^{(1)}, x_2^{(1)})^T$, with

$$x_1^{(1)} = x_2^{(1)} = \frac{\big(x_1^{(0)}\big)^2 + \big(x_2^{(0)}\big)^2 + 2}{2\big(x_1^{(0)} + x_2^{(0)}\big)}.$$

Prove convergence and deduce that the iteration converges to $(1, 1)^T$ if $x_1^{(0)} + x_2^{(0)}$ is positive, and, if $x_1^{(0)} + x_2^{(0)}$ is negative, the iteration converges to the other solution.

Solution: the Jacobian of $\mathbf{f}$ is

$$J = \begin{bmatrix} 2x_1 & 2x_2 \\ 1 & -1 \end{bmatrix}$$

and

$$\mathbf{x}^{(1)} = \mathbf{x}^{(0)} - \boldsymbol{\Delta}\mathbf{x}, \qquad J(\mathbf{x}^{(0)})\,\boldsymbol{\Delta}\mathbf{x} = \mathbf{f}(\mathbf{x}^0).$$

Solving the linear system and inserting in the first equation we get the expression for $\mathbf{x}^{(1)}$ and

$$x_1^{(1)} = x_2^{(1)} = \frac{\big(x_1^{(0)}\big)^2 + \big(x_2^{(0)}\big)^2 + 2}{2\big(x_1^{(0)} + x_2^{(0)}\big)}.$$

Since in this expression the numerator is always positive, then if $x_1^{(0)} + x_2^{(0)}$ is positive so are all $x_1^{(k)} = x_2^{(k)}$ for all k. Similarly if $x_1^{(0)} + x_2^{(0)}$ is negative. The case when $x_1^{(0)} + x_2^{(0)} = 0$ corresponds to the Jacobian $J(\mathbf{x}^{(0)})$ being not invertible.

We next prove convergence for the case $x_1^{(0)} + x_2^{(0)}$ positive, the negative case is analogous. We simplify the notation to $x^k := x_1^{(k)} = x_2^{(k)}$, $\{x^k\}_k$ for $k \geq 1$, and $\{x^k\}_k$ is the sequence

$$\frac{1}{2}\left(x^k + \frac{1}{x^k}\right).$$

If $x_1^{(0)} + x_2^{(0)} > 0$ then $x^1 > 0$ and we have three cases

$$x^1 = \begin{cases} 0 < x^1 < 1, \\ x^1 = 1, \\ 1 < x^1. \end{cases}$$

In the last case

$$1 < x^{k+1} = \frac{1}{2}\left(x^k + \frac{1}{x^k}\right) < x^k$$

for $k \geq 1$, a monotonically decreasing sequence tending to 1.[1] The middle case occurs when $x_1^0 = x_2^0 = 1$ and equal to the solution of the nonlinear system. In the first case we have

$$1 < x^2 = \frac{1}{2}\left(x^1 + \frac{1}{x^1}\right)$$

and we end up in the third case for $k \geq 2$.

b) Prove that convergence is quadratic.

Solution: Consider the error $e^k = 1 - x^k$, then

$$e^{k+1} = 1 - \frac{1}{2}\left(x^k + \frac{1}{x^k}\right) = \frac{1 - x^k}{2} - \frac{1}{2}\frac{(1 - x^k)}{x^k} = e^k \frac{1}{2}(1 - \frac{1}{x^k}) = -\frac{1}{2x^k}(e^k)^2.$$

[1] In this case x^1 is bigger than 1, so $\frac{1}{x^1} < 1$ and $\frac{1}{2}\left(x^1 + \frac{1}{x^1}\right)$ is an average between $x^1 > 1$ and $\frac{1}{x^1} < 1$, so it must be $\frac{1}{2}\left(x^1 + \frac{1}{x^1}\right) < x^1$. On the other hand, $|1 - \frac{1}{x^1}| < |x^1 - 1|$ so $1 < \frac{1}{2}\left(x^1 + \frac{1}{x^1}\right)$. This reasoning remains valid for $k \geq 2$.

So

$$\lim_{k\to\infty} \frac{|e^{k+1}|}{(|e^k|)^2} = \lim_{k\to\infty} \frac{1}{2|x^k|} = \frac{1}{2},$$

and the convergence is therefore quadratic.

Problem 2

Preliminaries

Recall that the spectral radius of a $N \times N$ real or complex matrix A is

$$\rho(A) := \max_{\lambda\in\sigma(A)} |\lambda| \quad \text{and} \quad \sigma(A) = \{\lambda \in \mathbf{C} \quad | \quad \lambda \quad \text{eigenvalue} \quad \text{of} \quad A\}.$$

Assuming $A = M - N$, with M invertible, an iterative method for linear systems $Ax = b$ can be formulated as follows:

$$x^{(n+1)} = M^{-1}Nx^{(n)} + M^{-1}b,$$

and by defining the **iteration matrix** $C := M^{-1}N$ and $g := M^{-1}b$ we can write:

$$x^{(n+1)} = Cx^{(n)} + g. \tag{1}$$

Theorem 1 *The iterative method (1) converges, independently on the choice of initial guess $x^{(0)}$, if and only if $\rho(C) < 1$.*

We will now prove a part of the lemma that we have used to prove Theorem 1.

Exercise

Let $\rho(A)$ be the spectral radius of A. Prove that for all operator norms[2] $\|\cdot\|$,

$$\rho(A) \le \|A\|.$$

Solution: Consider v to be the eigenvector associated with the eigenvalue μ such that $\rho(C) = \max_{\lambda\in\sigma(C)} |\lambda| = |\mu|$. Then for any vector norm $\|\cdot\|$, we consider the corresponding operator norm (subordinate matrix norm) of C and have

$$\|C\| := \max_{x\neq 0,\, x\in\mathbb{R}^N} \frac{\|Cx\|}{\|x\|} \ge \frac{\|Cv\|}{\|v\|} = |\mu| = \rho(C).$$

[2]subordinate matrix norms

Problem 3

Recall that when M is chosen to be the diagonal part of A, (1) is the so called **Jacobi method**; and when M is chosen to be the lower triangular part of A (diagonal included), (1) is the so called **Gauss-Seidel method**.

Exercise

Consider the matrix

$$A := \begin{bmatrix} 1 & -1 \\ -1 & 2 \end{bmatrix},$$

we want to analyse the convergence of the Jacobi and the Gauss-Seidel method for this matrix.

Guidelines:

- Compute the iteration matrices for the Jacobi and the Gauss-Seidel method.
- Compute the spectral radii of these matrices.
- Use Theorem 1 to conclude if the methods converge or not.

Solution: the solution of this exercise is in the note on Linear algebra (part 1 page 8).