Master in Mathematics and Applications - Técnico, Lisbon
Numerical Functional Analysis and Optimization - Fall Semester 2016

Exam (Part I) – January 19th, 2017 – Solutions

1. Let $V = C([0,1])$ and consider the integral operator $T : V \to V$ defined by

$$Tu(t) := \int_0^1 \cos(\lambda\, u(s))\, ds + f(t)\,, \tag{1}$$

where $\lambda \in \mathbb{R} \setminus \{0\}$ and $f \in V$.

a) Show that T is Fréchet differentiable and caracterize its Fréchet derivative $T'(u)$. **[1.0]**

Solution: Let $h \in V$, with $\|h\|_V \ll 1$. Given that

$$\cos(\lambda\,(u+h)) = \cos(\lambda\, u) - \lambda\, \sin(\lambda\, u)\, h - \frac{1}{2}\lambda^2 \cos(\lambda\, u)\, h^2 + \mathcal{O}(\|h\|_V^3)\,,$$

we see that

$$T(u+h) - Tu - T'(u)h = \mathcal{O}(\|h\|_V^2)\,,$$

where

$$T'(u)h = -\lambda \int_0^1 \sin(\lambda\, u(s))\, h(s)\, ds\,. \tag{2}$$

The operator T is thus Fréchet differentiable and its Fréchet derivative is caracterized by (2).

b) Consider the integral equation

$$u(t) - \int_0^1 \cos(\lambda\, u(s))\, ds = f(t)\,, \quad 0 \le t \le 1\,. \tag{3}$$

Show that there exists $\lambda_0 > 0$ such that for $|\lambda| \le \lambda_0$, equation (3) admits one and only one solution $u \in V$. **[1.5]**

Solution: Given that,

$$\|T'(u)\, h\|_V = \max_{t\in[0,1]} \left|\lambda \int_0^1 \sin(\lambda\, u(s))\, h(s)\, ds\ \right| \le |\lambda| \int_0^1 |\sin(\lambda\, u(s))|\, ds\, \|h\|_V \le |\lambda|\, \|h\|_V\,,$$

we have

$$\|T'(u)\|_{\mathcal{L}(V)} = \sup_{h\in V\setminus\{0\}} \frac{\|T'(u)\, h\|_V}{\|h\|_V} \le |\lambda| \qquad \forall\, u \in V\,.$$

Thus, if $\lambda \in \mathbb{R} \setminus \{0\}$ is such that $|\lambda| \le \lambda_0 < 1$, then $\|T'(u)\, h\|_V < 1$ for all $u \in V$, that is, the operator T is a contraction in V.

Moreover, $T : V \to V$ and $V = C[0,1]$ is a closed and non-empty set, thus by Banach's Fixed Point Theorem, the operator T admits a unique fixed point $u \in V$ provided $|\lambda| \le \lambda_0 < 1$. In other words, there exists $\lambda_0 \in (0,1)$ such that if $|\lambda| \le \lambda_0$ then the operator equation $u = T(u)$ admits a unique solution $u \in V$.

c) Let $f(t) = t$. Approximate the solution of the integral equation (3) by the fixed point method. Consider $\lambda = 0.5$, $u_0(t) = 0$ and compute the first three iterates. Derive an upper bound for the error

$$\|u - u_3\|_V = \max_{t\in[0,1]} |u(t) - u_3(t)|\,.$$ **[1.5]**

Solution: The fixed point method reads as

$$u_{n+1}(t) = \int_0^1 \cos(0.5\, u(s))\, ds + t\,, \quad n = 0, 1, \dots\,, \qquad u_0(t) = 0\,.$$

It follows that

$$u_1(t) = \textstyle\int_0^1 ds + t = 1 + t\,, \qquad u_2(t) = \int_0^1 \cos(0.5\,(1+s)\, ds + t = 2\,(\,\sin 1 - \sin 0.5\,) + t = 0.724091 + t\,,$$

$$u_3(t) = \textstyle\int_0^1 \cos(0.5\,(0.724091 + s)\, ds + t = 2\left(\,\sin \frac{1+0.724091}{2} - \sin \frac{0.724091}{2}\,\right) + t = 0.809975 + t\,.$$

Recalling the error estimate

$$\|u - u_3\|_V \le \frac{\alpha}{1-\alpha}\, \|u_3 - u_2\|_V\,,$$

where $\alpha \in (0,1)$ is the Lipschitz constant of T. It holds

$$\alpha \le |\lambda| = 0.5\,, \quad \|u_3 - u_2\|_V = 0.085884 \qquad \Rightarrow \qquad \|u - u_3\|_V \le 0.085884\,.$$

2. Let $u, v \in \mathbb{R}^N \setminus \{0\}$ be two column vectors. Show that the matrix $I + uv^T$ is invertible if and only if $1 + v^T u \neq 0$. **[1.5]**

Solution: The eigenvalues of matrix $I + uv^T$ are 1 and $1 + v^T u$. In fact,

$$(I + uv^T)\, u = (1 + v^T u)\, u\,, \qquad (I + uv^T)\, w_j = w_j\,, \quad j = 1, 2, \dots N-1\,,$$

where $w_j, j = 1, 2, \dots N-1$ are such that $(v, w_j) = 0$. Note that there exist $N-1$ linearly independent vectors orthogonal to a given vector $v \in \mathbb{R}^N \setminus \{0\}$. Since all eigenvalues of $I + uv^T$ are non-zero, the matrix $I + uv^T$ is non-singular.

3. Consider the vector field $F : \mathbb{R}^2 \to \mathbb{R}^2$ defined by

$$F(x) = \begin{bmatrix} x_1 + x_2 - 3 \\ x_1 + 2x_2 - 9 \end{bmatrix}.$$

a) Show that the Newton's method converges to the exact solution $x_* = [-3\ 6]^T$ of equation $F(x) = 0$ in one iteration for any initial approximation. **[1.0]**

Solution: Let $F(x) = Ax - b$, where

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad b = \begin{bmatrix} 3 \\ 9 \end{bmatrix}.$$

It is easy to see that $Ax_* - b = 0$ and that A is a non-singular matrix so $x_* = A^{-1}b = [-3\ 6]^T$ is the unique solution of equation $F(x) = 0$. Newton's method reads as

$$x^{(n+1)} = x^{(n)} - J_F^{-1}(x^{n)})\, F(x^{(n)})\,, \qquad n = 0, 1, \dots\,.$$

Given that $J_F(x) = A$, it follows that

$$x^{(1)} = x^{(0)} - A^{-1}\,(A\, x^{(0)} - b) = A^{-1}\, b\,,$$

independently of the initial guess $x^{(0)} \in \mathbb{R}^2$.

b) Approximate the solution of equation $F(x) = 0$ by Broyden's method. Consider

$$x_0 = [-1\ 5]^T\,, \qquad B_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

and compute the first three iterates. Explain the result. **[1.5]**

<u>Solution:</u> In Broyden's method, we need to, given B_0 and x_0, solve the linear system $B_k s_k = -F(x_k)$ and compute

$$x_{k+1} = x_k + s_k\,, \qquad y_k = F(x_{k+1}) - F(x_k)\,, \qquad B_{k+1} = B_k + \frac{(y_k - B_k s_k)s_k^T}{s_k^T s_k}\,.$$

for $k = 0, 1, \ldots$. Given that $As_k = y_k$, we may write the Broyden's update formula as

$$B_{k+1} = B_k + \Big(A - B_k\Big)\frac{s_k s_k^T}{s_k^T s_k}\,, \qquad \text{where} \qquad A = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}.$$

It follows that

$$F(x_0) = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \qquad s_0 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}, \qquad x_1 = \begin{bmatrix} -2 \\ 5 \end{bmatrix}.$$

$$F(x_1) = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \qquad s_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}, \qquad x_2 = \begin{bmatrix} -2 \\ 6 \end{bmatrix}.$$

$$F(x_2) = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \qquad s_2 = \begin{bmatrix} -1 \\ 0 \end{bmatrix}, \qquad x_3 = \begin{bmatrix} -3 \\ 6 \end{bmatrix} = x_*\,.$$

The third iterate coincides with the exact solution. In fact, Broyden's method converges to the exact solution of a linear system of N equations at most in $2N$ iterations.

4. Let K be a non-empty, closed, convex subset of a Hilbert space H, equipped with the inner product $(\cdot,\cdot)$, and assume that $f \in H$. Show that $\phi^* \in K$ is the best approximation of f in K, with respect to the induced norm, if and only if

$$(f - \phi^*, \phi - \phi^*) \leq 0 \qquad \forall\, \phi \in K\,. \qquad \textbf{[2.0]}$$

<u>Solution:</u> Let ϕ^* be a best approximation of $f \in H$ in K and ϕ an arbitrary element in K. Since K is convex, we have

$$\phi^* + \lambda\,(\phi - \phi^*) \in K \qquad \forall\, \lambda \in [0,1]\,.$$

This means that the real-valued and differentiable function $\varphi : [0,1] \to \mathbb{R}$ defined by

$$\varphi(\lambda) = \|f - (\phi^* + \lambda\,(\phi - \phi^*))\|^2$$

has a minimum at $\lambda = 0$. Therefore $\varphi'(0) \geq 0$. Observing that

$$\varphi(\lambda) = \|f - \phi^*\|^2 - 2\,\lambda\,(f - \phi^*, \phi - \phi^*) + \lambda^2\|\phi - \phi^*\|^2\,,$$

we obtain

$$\varphi'(0) = -2\,(f-\phi^*, \phi-\phi^*) \geq 0 \qquad \Leftrightarrow \qquad (f-\phi^*, \phi-\phi^*) \leq 0\,.$$

On the other hand, if $(f-\phi^*, \phi-\phi^*) \leq 0$, then

$$\|f-\phi\|^2 = \|f-\phi^*+\phi^*-\phi\|^2 = \|f-\phi^*\|^2 + 2\,(f-\phi^*, \phi^*-\phi) + \|\phi^*-\phi\|^2$$

$$\geq \|f-\phi^*\|^2 + \|\phi^*-\phi\|^2 \geq \|f-\phi^*\|^2 \qquad \forall\,\phi \in K\,,$$

that is $\phi^* \in K$ is, by definition, a best approximation of $f \in H$ in K with respect to the induced norm $\|\cdot\| = \sqrt{(\cdot,\cdot)}$.

Exam (Part II) – January 19th, 2017 – Solutions

1. Consider the unconstrained minimization problem

$$\min_{x \in S} \; f(x)\,,$$

where $f : S \to \mathbb{R}$ is a convex function, $S \subset \mathbb{R}^N$ a convex set and $x_* \in S$ a local minimizer of f. Show that x_* is a global solution of the problem. **[2.0]**

Solution: Assume, for the sake of contradiction, that x_* is not a global minimizer, that is, there exists $y \in S$, say, for which $f(y) < f(x_*)$. Given that S is a convex set, the convex combination $z(\lambda) = \lambda x_* + (1-\lambda)y$ of y and x_* is an element in S for all $\lambda \in [0,1]$. On the other, from the convexity of f it follows that

$$f(z(\lambda)) = f(\lambda x_* + (1-\lambda)y) \leq \lambda f(x_*) + (1-\lambda)f(y) < \lambda f(x_*) + (1-\lambda)f(x_*) = f(x_*)\,, \quad \forall\, \lambda \in (0,1)\,.$$

But this contradicts the fact that x_* is a local minimizer since $z(\lambda) \to x_*$ when $\lambda \to 1$.

2. Consider the unconstrained minimization problem

$$\min_{x \in \mathbb{R}^N} f(x)\,, \tag{4}$$

where $f(x) = \frac{1}{2}\, x^T A x$, with $A \in \mathbb{R}^{N \times N}$ symmetric and positive definite..

a) Show that the numerical approximation of problem (5) by the method of steepest descent corresponds to: Given $x_0 \in \mathbb{R}^N$ compute

$$x_{k+1} = x_k - \frac{x_k^T A^2 x_k}{x_k^T A^3 x_k}\; A x_k\,, \qquad k = 0, 1, \ldots \qquad \mathbf{[1.5]}$$

Solution: For the objective function $f(x) = \frac{1}{2} x^T A x$, it holds $\nabla f(x) = Ax$ and $H_f(x) = A$, since A is a symmetric matrix. The method of steepest descent thus reduces to

$$x_{k+1} = x_k - \alpha_k\, \nabla f(x_k) = x_k - \alpha_k A x_k\,, \tag{5}$$

where α_k solves the one-dimensional minimization problem

$$\min_{\alpha_k > 0} f(x_k - \alpha_k A x_k)\,.$$

To solve this problem exactly, we define

$$g(\alpha_k) := f(x_k - \alpha_k A x_k)\,,$$

and compute

$$g'(\alpha_k) = -(Ax_k)^T \nabla f(x_k - \alpha_k A x_k) = -(Ax_k)^T A(x_k - \alpha_k A x_k) = -x_k^T A^2 x_k + \alpha_k x_k^T A^3 x_k\,.$$

It follows that

$$g'(\alpha_k) = 0 \quad \Leftrightarrow \quad \alpha_k = \frac{x_k^T A^2 x_k}{x_k^T A^3 x_k}\,.$$

Note that $g''(\alpha_k) = x_k^T A^3 x_k$ and that A^2 and A^3 are positive definite matrices (given that A is the positive definite), so that α_k is the global positive minimizer of $g(\alpha_k)$. Substituting the value of α_k to (5), the method of steepest descent becomes

$$x_{k+1} = x_k - \frac{x_k^T A^2 x_k}{x_k^T A^3 x_k} A x_k, \quad k = 0, 1, \ldots$$

b) Consider the problem

$$\min_{x \in \mathbb{R}^N} F(x)\,, \tag{6}$$

where $F(x) = \frac{1}{2}\|\nabla f(x)\|_2^2$. Prove that in this case the method of steepst descent reduces to: Given $x_0 \in \mathbb{R}^N$, compute

$$x_{k+1} = x_k - \frac{x_k^T A^4 x_k}{x_k^T A^6 x_k} A^2 x_k\,, \qquad k = 0, 1, \ldots$$ [**1.0**]

<u>Solution:</u> The objective function F can be written as

$$F(x) = \frac{1}{2}\|\nabla f(x)\|_2^2 = \frac{1}{2}\|Ax\|_2^2 = \frac{1}{2}(Ax)^T A\, x = \frac{1}{2}\, x^T A^T A\, x = \frac{1}{2}\, x^T A^2 x\,,$$

where we have used the symmetry of A. This shows that minimizing F corresponds to minimizing f with the matrix A replaced with A^2. Thus the method of steepest descent for minimizing the objective function F can be written as

$$x_{k+1} = x_k - \frac{x_k^T A^4 x_k}{x_k^T A^6 x_k} A^2 x_k, \quad k = 0, 1, \ldots.$$

c) Show that the sequence $\{f(x_k)\}$ converges to $f(x_*)$ and the sequence $\{F(x_k)\}$ converges to $F(x_*)$ when $k \to \infty$. Show also that the convergence of the sequence $\{f(x_k)\}$ is faster that that of the sequence $\{F(x_k)\}$. **[2.0]**

<u>Solution:</u> Defining the error $e_{k+1} = f(x_{k+1}) - f(x_*)$, we have the estimate

$$|e_{k+1}| \le \left(\frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1}\right)^2 |e_k|,$$

where $0 < \lambda_1 \le \ldots \le \lambda_N$ are the eigenvalues of the (symmetric and positive definite) Hessian matrix $H_f(x_*) = A$, cf. formulas. For the objective function F, the error estimate can be written as

$$|E_{k+1}| \le \left(\frac{\mu_N - \mu_1}{\mu_N + \mu_1}\right)^2 |E_k|,$$

where $E_{k+1} = F(x_{k+1}) - F(x_*)$ and $0 < \mu_1 \le \ldots \le \mu_N$ are the eigenvalues of the Hessian matrix $H_F(x_*) = A^2$.

Note that if $x \in \mathbb{R}^N \setminus \{0\}$ is an eigenvector associated with the eigenvalue λ of matrix A, it holds

$$A^2 x = A(Ax) = A\lambda x = \lambda A x = \lambda^2 x\,,$$

that is, the eigenvalues of matrix A^2 are $\mu_k = \lambda_k^2,\ k = 1, \ldots, N$. If $\lambda_1 < \lambda_N$, it thus holds

$$0 < \frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} = 1 - 2\,\frac{\lambda_1}{\lambda_N + \lambda_1} < 1\,, \qquad 0 < \frac{\mu_N - \mu_1}{\mu_N + \mu_1} = 1 - 2\,\frac{\lambda_1^2}{\lambda_N^2 + \lambda_1^2} < 1\,,$$

so that

$$|e_{k+1}| \leq \beta^{k+1} |e_0| , \qquad |E_{k+1}| \leq \gamma^{k+1} |E_0| ,$$

with some $\beta, \gamma \in (0,1)$. The sequence $\{f(x_k)\}$ thus converges to $f(x_*)$ and the sequence $\{F(x_k)\}$ converges to $F(x_*)$.

To show that the first method converges more rapidly, it suffices to observe that if $\lambda_1 \neq \lambda_N$ then

$$\frac{\lambda_N - \lambda_1}{\lambda_N + \lambda_1} < \frac{\mu_N - \mu_1}{\mu_N + \mu_1} = \frac{\lambda_N^2 - \lambda_1^2}{\lambda_N^2 + \lambda_1^2} \iff \lambda_N^2 + \lambda_1^2 < (\lambda_N^2 + \lambda_1^2)^2 \iff 2\lambda_1\lambda_N > 0 ,$$

where the last inequality holds since $\lambda_j > 0 \; \forall j$.

Obs.: When $\lambda_1 = \lambda_N$, i.e. $\lambda_1 = \lambda_2 = \ldots = \lambda_N$, we have $|e_1| = |E_1| = 0$, that is, both methods attain the exact solution x_* in a single iteration. Note also that both methods converge to the global minimizer $x_* = 0$.

3. Consider the following constrained minimization problem

$$\min_{x \in \mathbb{R}^N} x^T A x \qquad \text{subject to} \qquad x^T x = 1 , \tag{7}$$

where $A \in \mathbb{R}^{N \times N}$ is a symmetric matrix with N distinct eigenvalues $\lambda_1 < \lambda_2 < \ldots < \lambda_N$.

a) Write down the KKT conditions for problem (7). Determine all KKT points (x_*, λ^*). Show that the constraint qualification LICQ is valid at the stationary points x_*. **[2.0]**

Solution: The constrained minimization problem reads as

$$\min_{x \in \mathbb{R}^N} f(x) \qquad \text{subject to} \qquad c_1(x) = 0 , \tag{8}$$

where $f(x) = x^T A x$ and $c_1(x) = x^T x - 1$. The Lagrangean function associated with problem (8) is thus $\mathcal{L}(x, \lambda) = f(x) - \lambda\, c_1(x)$. The KKT conditions hold at (x_*, λ_*) if

$$\nabla_x \mathcal{L}(x_*, \lambda^*) = 0 , \qquad c_1(x_*) = 0 ,$$

that is, if $2\,A\,x_* - 2\,\lambda^* x_* = 0$ and $x_*^T x_* = 1$. The Lagrange multiplier λ^* satisfying the KKT conditions thus corresponds to one of the N distinct eigenvalues $\lambda_j, j = 1, \ldots, N$, of matrix A, and the stationary points x_* are the (normalized) eigenvectors w_j associated with λ_j's, i.e.

$$A\,x_* = \lambda^* x_* , \qquad \|x_*\|_2 = 1 , \quad \text{where} \quad (x_*, \lambda^*) = (w_j, \lambda_j) , \quad j = 1, \ldots, N .$$

The constraint qualification LICQ is valid if $\nabla c_1(x_*) \neq 0$. Now, $\nabla c_1(x_*) = 2\,x_* \neq 0$ since $\|x_*\|_2 = 1$; thus LICQ holds at each of the N stationary points.

b) Determine all local and global solutions of problem (7). **[1.5]**

Solution: The local solutions (minimizers) of problem (7) must satisfy the second-order necessary conditions:

$$w^T \nabla_x^2 \mathcal{L}(x_*, \lambda^*)\, w \geq 0 \quad \forall w \in F_2(x_*, \lambda^*) ,$$

where $F_2(x_*, \lambda^*) = \{w \in \mathbb{R}^N \,|\, \nabla c_1(x_*)^T w = 0\}$. Given that A is symmetric, $\nabla_x^2 \mathcal{L}(x_*, \lambda^*) = 2\,A - 2\lambda^* I$, and thus

$$w^T \nabla_x^2 \mathcal{L}(x_*, \lambda^*)\, w = 2\left(w^T A\, w - \lambda^* w^T w \right) . \tag{9}$$

Now, fix $j \in \{2, \ldots, N\}$ and consider the KKT point $(x_*, \lambda^*) = (w_j, \lambda_j)$, where w_j is the eigenvector associated with λ_j. The matrix A is symmetric, thus its eigenvectors constitute an orthogonal basis in $\mathbb{R}^N$. Thus, the critical cone $F_2(x_*, \lambda^*)$, spanned by vectors w orthogonal to w_j, is composed of the $N-1$ eigenvectors $w_k, k \neq j$ of A, including, in particular, w_1. Hence, testing the condition (9) at $(x_*, \lambda^*) = (w_j, \lambda_j)$, with $w = w_1 \in F_2(x_*, \lambda^*)$, we get

$$w^T \nabla_x^2 \mathcal{L}(x_*, \lambda^*)\, w = 2\Big(\, w_1^T A\, w_1 - \lambda_j w_1^T w_1 \Big) = 2(\lambda_1 - \lambda_j)\|w_1\|_2^2 < 0 \quad \forall j \neq 1\,,$$

since λ_1, the eigenvalue associated with the eigenvector w_1, is the smallest eigenvalue and the eigenvalues are distinct. The second-order necessary condition of optimality is thus violated and we conclude that the $N-1$ stationary points $(x_*, \lambda^*) = (w_j, \lambda_j), j = 2, \ldots, N$, are not local solutions.

On the other hand, when $(x_*, \lambda^*) = (w_1, \lambda_1)$, we have $F_2(x_*, \lambda^*) = \text{span}\{w_2, w_3, \ldots, w_N\}$ and

$$w^T \nabla_x^2 \mathcal{L}(x_*, \lambda^*)\, w = 2\Big(\, w_j^T A\, w_j - \lambda_1 w_j^T w_j \Big) = 2(\lambda_j - \lambda_1)\|w_j\|_2^2 > 0 \quad \forall\, w = w_j\,, \quad j = 2, 3, \ldots, N\,.$$

In other words, the second-order sufficient condition is valid and, therefore, $(x_*, \lambda^*) = (w_1, \lambda_1)$ is a local minimizer.

The point $(x_*, \lambda^*) = (w_1, \lambda_1)$ is also a global solution since by the Rayleigh quotient

$$f(x_*) = x_*^T\, A x_* = \lambda_*\, x_*^T x_* = \lambda_1 \leq x^T A x = f(x) \quad \forall\, x \in \mathbb{R}^N\,, \quad \text{with} \;\; x^T x = 1\,,$$

or because, by Weierstrass' theorem, the objective function must attain its global minimum (and maximum) in the feasible set S since $f : S \to \mathbb{R}^N$ is continuous and S is compact. The solution is $(x_*, \lambda^*) = (w_1, \lambda_1)$ is unique because the LICQ condition is valid at each point of the feasible region so that all possible solutions must satisfy the first-order necessary conditions (KKT conditions).