MATHEMATICAL FOUNDATIONS OF SIGNAL PROCESSING Fall 2019

Adam Scholefield, Mihailo Kolundžija, Reza Parhizkar
Teaching assistants: Michalina Pacholska

Solution to Homework 2 (Graded)

Friday, November 29, 2019

Exercise 1. Quick Review of Chapter 2

(i) In the following, we have to check that the system is linear ($A(\alpha x + \beta y) = \alpha Ax + \beta Ay$ for any scalars α, β) and shift invariant ($A(x(t-\tau)) = (Ax)(t-\tau)$):

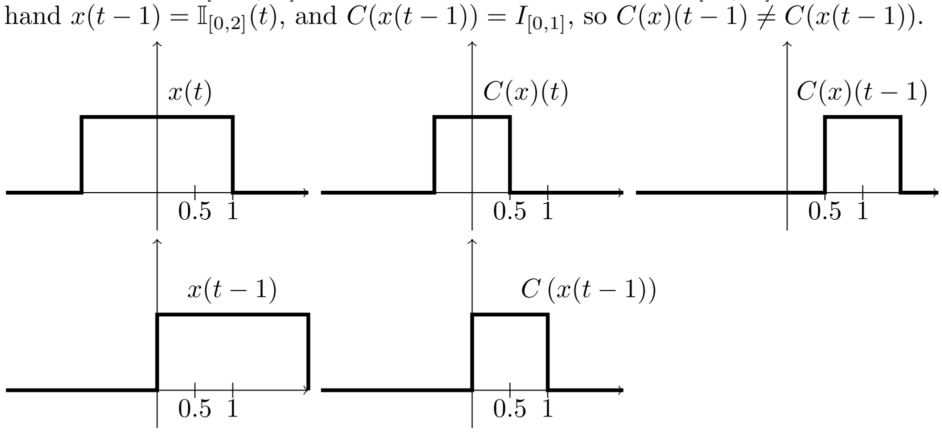
- We can use the fact that convolution is linear and shift invariant. But we can also show this directly:

$$\begin{aligned} A(\alpha x + \beta y)(k) &= \sum_n (\alpha x[n] + \beta y[n]) h_A[k-n] \\ &= \sum_n \alpha x[n] h_A[k-n] + \sum_n \beta y[n] h_A[k-n] \\ &= \alpha \sum_n x[n] h_A[k-n] + \beta \sum_n y[n] h_A[k-n] \\ &= \alpha Ax(k) + \beta Ay(k) \end{aligned}$$

And similarly:

$$\begin{aligned} A(x[n-m]) &= \sum_n x[n-m] h_A[k-n] \\ &= \sum_n x[n-m] h_A[(k-m)-(n-m)] \\ &= Ax[k-m] \end{aligned}$$

- B is not linear, because for $\alpha = 0$ and any x we have $B(0 \cdot x)(t) = \operatorname{sinc}(t) \neq 0 = 0 \cdot B(x)(t)$
- C is not shift invariant. Consider indicator function $x = \mathbb{I}_{[-1,1]}$, depicted below. Then $C(x)(t) = \mathbb{I}_{[-0.5,0.5]}$, and consequently $C(x)(t-1) = \mathbb{I}_{[0.5,1.5]}$. On the other hand $x(t-1) = \mathbb{I}_{[0,2]}(t)$, and $C(x(t-1)) = I_{[0,1]}$, so $C(x)(t-1) \neq C(x(t-1))$.



- D is linear and shift invariant (from the definition of the derivative). The linearity of derivative should be clear, but for shift invariance, let $s = t - \tau$. Then $D(x)(s) = \frac{dx}{ds}(s) = \frac{dx}{dt}(t-\tau) = D(x(t-\tau))$.

(ii) Rewriting the adjoint definition using the inner product of $\mathcal{L}^2[a,b]$ gives

$$\langle h, \mathcal{T} f\rangle = \langle \mathcal{T}^* h, f\rangle \iff \int_a^b \overline{h(x)}(\mathcal{T} f)(x)dx = \int_a^b \overline{(\mathcal{T}^* h)(r)} f(r)dr,$$

Substituting the kernel representations for $\mathcal{T}$ and $\mathcal{T}^*$, we obtain

$$\begin{aligned} &\int_a^b \overline{h(x)}\big(\int_a^b T(x,y)f(y)dy\big)dx = \int_a^b \overline{\big(\int_a^b T^*(r,\rho)h(\rho)d\rho\big)} f(r)dr \\ \implies &\int_a^b \int_a^b \overline{h(x)} T(x,y) f(y)dydx = \int_a^b \int_a^b \overline{h(\rho)}\ \overline{T^*(r,\rho)} f(r)drd\rho \end{aligned}$$

which can be written as

$$\int_a^b \int_a^b \overline{h(x)}\big(T(x,y) - \overline{T^*(y,x)}\big) f(y)dydx = 0$$

which leads to $T^*(y,x) = \overline{T(x,y)}$.

(iii) Again, from the definition of adjoint operator, we are looking for an operator C^* such that $\langle y, Cx\rangle = \langle C^* y, x\rangle$. Expanding the inner product we get:

$$\begin{aligned} \langle y, Cx\rangle &= \int_{-\infty}^{\infty} y^*(t)x(2t)dt \\ &= \int_{-\infty}^{\infty} \frac{1}{2} y^*\left(\frac{s}{2}\right) x(s)ds, \end{aligned}$$

where the second equality comes from the change of variables. We get that the adjoint is $C^* x(t) = \frac{1}{2} x\left(\frac{t}{2}\right)$.

(iv) Yet again, from the definition of adjoint operator, we are looking for an operator D^* such that $\langle y, Dx\rangle = \langle D^* y, x\rangle$., where $x, y \in \mathcal{H}$. Expanding the inner product we get:

$$\langle y, Dx\rangle = \int_{-\infty}^{\infty} y^*(t)\frac{dx}{dt}(t)dt = \int_{-a}^{b} y^*(t)\frac{dx}{dt}(t)dt,$$

were a, b are some constants dependent on x, y. Integrating by parts we get that:

$$\begin{aligned} \langle y, Dx\rangle &= [x(t)y(t)]_a^b - \int_{-a}^{b} \frac{dy^*}{dt}(t)x(t)dt \\ &= \int_{-a}^{b} \frac{dy^*}{dt}(t)x(t)dt \\ &= \int_{-\infty}^{\infty} \frac{dy^*}{dt}(t)x(t)dt, \end{aligned}$$

so we get $D^* x = -\frac{dx}{dt} = -Dx$.

Solution 1. LCMV AND GSC DERIVATION

(i) Any x' for which $A\,x' = b$ can be written as $x' = x + z$, where $z \in \mathcal{N}(A)$ and x is the minimum norm solution of our optimization problem. The minimum norm solution x is given by projecting the zero vector onto the constraint hyperplane $A\,x = b$. Vector x is thus orthogonal to $\mathcal{N}(A)$, i.e. $x \in \mathcal{R}(A^*)$. In other words, we can write $x = A^*\,y$, which brings us to

$$A\,A^*\,y = b \text{ and } x = A^*\,y\,.$$

(ii) If A is of full rank, $A\,A^*$ is invertible and $x = A^*(A\,A^*)^{-1}\,b$.

(iii) Since R_x is a positive definite covariance matrix, its eigenvalues are all positive and it can thus be written as $R_x = V\,\Lambda_s\Lambda_s\,V^*$, where V is an $M \times M$ matrix with orthonormal eigenvectors of R_x, while Λ_s is a diagonal matrix with positive entries equal to square roots of eigenvalues of R_x.
We define an $M \times M$ matrix B as $B = \Lambda_s\,V^*$. Matrix B contains scaled eigenvectors of R_x as columns and is of full rank; we also have $R_x = B^*\,B$. The inverse of B is given by $B^{-1} = V\,\Lambda_s^{-1}$, where Λ_s^{-1} is a diagonal matrix whose diagonal elements are reciprocals of elements of Λ_s.
The original optimization problem can then be rewritten in the following way:

$$h = \underset{h}{\operatorname{argmin}}\; h^*\,B^*\,B\,h \quad \text{subject to} \quad C^*\,h = f\,.$$

By noticing that any vector h can be written as $h = B^{-1}\,g$ (remember that B is a basis), the optimization problem can be rephrased as finding the vector g such that $g = B\,h$ and

$$g = \underset{g}{\operatorname{argmin}}\; g^*\,g \quad \text{subject to} \quad C^*\,B^{-1}\,g = f\,.$$

This we recognize to be the optimization problem from the first part of the exercise, where the matrix $C^*\,B^{-1}$ plays the role of the constraint matrix A. The fact that C and B^{-1} are of full rank then makes it possible to express the solution as

$$h = B^{-1}(B^*)^{-1}\,C(C\,B^{-1}(B^*)^{-1}\,C^*)^{-1}\,f\,. \tag{1}$$

Since $R_x = B^*\,B$, we can replace $B^{-1}(B^*)^{-1}$ with R_x^{-1} to arrive at

$$h = R_x^{-1}\,C\left(C^*R_x^{-1}C\right)^{-1} f\,.$$

(iv) We denote by $\mathcal{L}(h_n)$ the cost function that is minimized, namely $\mathcal{L}(h_n) = (h_0 - C_n\,h_n)^*\,R_x\,(h_0 - C_n\,h_n)$. Since $\mathcal{L}(h_n)$ is a quadratic form, its minimum is given by its stationary point

$$\tfrac{\partial \mathcal{L}}{\partial h_n} = 2\,C_n^*\,R_x\,C_n\,h_n - 2\,C_n^*\,h_0 = 0\,.$$

Since C_n and R_x are of full rank, the matrix $C_n^*\,R_x\,C_n$ is invertible, so we have

$$h_n = (C_n^*\,R_x\,C_n)^{-1}C_n^*\,R_x\,h_0\,.$$