Solutions to Homework Set One
ECE 175
Electrical and Computer Engineering
University of California San Diego

Nuno Vasconcelos

1. a)

i) the column space of $\mathbf{A}$ is the space spanned by the columns, i.e. the set of vectors of the form

$$\mathbf{v} = \alpha\mathbf{a} + \beta\mathbf{b}$$

where $\mathbf{a} = (1, 0)^T$ and $\mathbf{b} = (1, 2)^T$, for any α and β.

ii) the row space of $\mathbf{A}$ is the space spanned by the rows, i.e. the set of vectors of the form

$$\mathbf{v} = \alpha\mathbf{c} + \beta\mathbf{d} \tag{1}$$

where $\mathbf{c} = (1, 1, 0)$ and $\mathbf{d} = (0, 2, 0)^T$, for any α and β.

iii) the null space of $\mathbf{A}$ is the space of vectors orthogonal to the row space. These are the vectors orthogonal to all rows, i.e. the set of vectors $\mathbf{v} = (v_1, v_2, v_3)^T$, such that $\mathbf{A}\mathbf{v} = 0$. This leads to

$$\begin{aligned} v_1 &= -v_2 \\ v_2 &= 0 \end{aligned}$$

i.e. $v_1 = v_2 = 0$. It follows that

$$\mathcal{N}(\mathbf{A}) = \{\mathbf{v}|\mathbf{v} = (0, 0, v_3)^T\},$$

iv) to find the rank we count the number of linearly independent rows or columns. Note that the only way to make the sum of (1) equal to zero is to make $\alpha = \beta = 0$. Hence, the rows are linearly independent, and since we only have these two, the rank is 2.

iv) the dimension of the column space is the same as the rank of the matrix, and therefore 2.

b) i) Note that the integral equations given in the problem are the dot-products between the functions 1, $\sin(mx)$ and $\cos(mx)$ for all possible values of m. Clearly, these functions are all orthogonal, since only the dot-product of the function with itself is non-zero. Hence the functions are linearly independent. They also span the space, since all $f(x)$ are linear combinations of these functions (this is how we define the space itself, see the expression of $f(x)$ in the statement of the problem). It follows that the functions are a basis of $\mathcal{F}$.

ii) The basis is orthogonal, but not orthonormal. For this, the norm of all basis functions (i.e. the dot-product of the function with itself) would have to be 1. We can see from the integral equations that this is not the case for $\sin(mx)$ and $\cos(mx)$ which have norm $\sqrt{\pi}$.

iii) The dimension is the number of function in the basis, which is ∞ in this case.

iv) The coordinates of the projection are given by the dot-product of the function with the basis functions. Let's consider the case of $\cos(mx)$. Using the integral equations we can write

$$\begin{aligned} < f(x), \cos(mx) > &= \frac{1}{2}a_0 < 1, \cos(nx) > + \sum_{n=1}^{\infty} a_n < \cos(mx), \cos(nx) > + \sum_{n=1}^{\infty} b_n < \cos(mx), \sin(nx) > \\ &= 0 + a_m\pi + 0 = a_m\pi. \end{aligned}$$

Similarly, we can show that

$$< f(x), \sin(mx) >= b_m\pi$$

and

$$< f(x), 1 >= a_0\pi$$

from which the projection vector is $\pi(a_0, a_1, ..., b_0, b_1, ...)$.

2. a) Since X and Y have values in $\{0, \ldots, n\}$ Z will have values in $\{0, \ldots, 2n\}$. Furthermore, for any such k

$$\begin{aligned} p_Z(k) &= Prob[Z = k] \\ &= Prob[X + Y = k] \\ &= \sum_{j=0}^{n} Prob[X + Y = k|Y = j]Prob[Y = j] \\ &= \sum_{j=0}^{n} Prob[X = k - j|Y = j]p_Y(j). \end{aligned}$$

Since X and Y are independent

$$Prob[X = k - j|Y = j] = Prob[X = k - j] = p_X(k - j)$$

and it follows that

$$p_Z(k) = \sum_{j=0}^{n} p_X(k - j)p_Y(j). \tag{2}$$

This is the convolution of p_X and p_Y.

b) In general, the result is that the sum of two independent random variables has a probability density function (pdf) which is the convolution of the pdfs of the two variables. This is a very neat result, that we can actually turn around and use to compute convolutions. In particular, if we have two functions $g_1(x)$ and $g_2(x)$ that satisfy the following two properties

1. $g_i(x) \geq 0, \forall x, i \in \{1, 2\}$
2. $\int g_i(x)dx = 1, \forall i \in \{1, 2\}$

we can always think of them as pdfs of independent random variables. If we denote by X_i the random variable associated with $g_i(x)$ and by $f(x)$ the result of the convolution that we are trying to determine

$$f = g_1 \star g_2$$

then $f(x)$ is the pdf of the random variable

$$Y = X_1 + X_2.$$

This trick is particularly useful for Gaussians, since the pdf of Y is extremely easy to determine in this case: the sum of two Gaussians is always a Gaussian, its expected value is the sum of the two expected values and, because the two Gaussian random variables being added are independent, the variance is the sum of the two variances (note that this only holds because the variables are independent). Hence,

$$f(x) = \mathcal{G}(x, \mu, \sigma) \tag{3}$$

where

$$\begin{aligned} \mu &= \mu_1 + \mu_2 \\ \sigma^2 &= \sigma_1^2 + \sigma_2^2. \end{aligned}$$

3. a) From $\mathbf{y} = \mathbf{A}\mathbf{x}$ it follows that

$$y_i = \sum_j a_{ij} x_j \tag{4}$$

and, from the linearity of the expected value,

$$E[y_i] = \sum_j a_{ij} E[x_j]. \tag{5}$$

Hence

$$\mu_{\mathbf{y}} = \mathbf{A}\mu_{\mathbf{x}}. \tag{6}$$

To compute $\mathbf{\Sigma_y}$ we note that

$$\begin{aligned} \mathbf{\Sigma_y} &= E[(\mathbf{y} - \mu_{\mathbf{y}})(\mathbf{y} - \mu_{\mathbf{y}})^T] \\ &= E[(\mathbf{A}\mathbf{x} - \mathbf{A}\mu_{\mathbf{x}})(\mathbf{A}\mathbf{x} - \mathbf{A}\mu_{\mathbf{x}})^T] \\ &= E[\mathbf{A}(\mathbf{x} - \mu_{\mathbf{x}})(\mathbf{x} - \mu_{\mathbf{x}})^T \mathbf{A}^T] \end{aligned}$$

Defining

$$\mathbf{\Gamma}(\mathbf{x}) = (\mathbf{x} - \mu_{\mathbf{x}})(\mathbf{x} - \mu_{\mathbf{x}})^T$$

and using the fact that the i, j^{th} element of $\mathbf{A\Gamma}(\mathbf{x})\mathbf{A}^T$ is

$$\left(\mathbf{A\Gamma}(\mathbf{x})\mathbf{A}^T\right)_{ij} = \sum_{kl} a_{ik} \left[\mathbf{\Gamma}(\mathbf{x})\right]_{kl} a_{jl}$$

we have

$$\begin{aligned} E\left[\left(\mathbf{A\Gamma}(\mathbf{x})\mathbf{A}^T\right)_{ij}\right] &= \sum_{kl} a_{ik} E\left\{\left[\mathbf{\Gamma}(\mathbf{x})\right]_{kl} a_{jl}\right\} \\ &= \sum_{kl} a_{ik} E\left\{\left[\mathbf{\Gamma}(\mathbf{x})\right]_{kl}\right\} a_{jl}. \end{aligned}$$

Hence

$$\begin{aligned} \mathbf{\Sigma_y} &= E\left[\mathbf{A\Gamma}(\mathbf{x})\mathbf{A}^T\right] \\ &= \mathbf{A} E\left[\mathbf{\Gamma}(\mathbf{x})\right] \mathbf{A}^T \\ &= \mathbf{A\Sigma_x A}^T. \end{aligned}$$

b) Consider any random variable $\mathbf{x}$ with covariance $\mathbf{\Sigma_x}$ and the transformation

$$y = \mathbf{a}^T \mathbf{x} \tag{7}$$

with $\mathbf{a}$ a column vector. Note that y is a scalar, but we can still use **a)** to obtain

$$\sigma_y^2 = \mathbf{a}^T \boldsymbol{\Sigma}_{\mathbf{x}} \mathbf{a}. \tag{8}$$

Since the variance of a random variable can never be negative we have

$$\mathbf{a}^T \boldsymbol{\Sigma}_{\mathbf{x}} \mathbf{a} \geq 0. \tag{9}$$

Note that, because we did not make any assumptions about $\mathbf{a}$, this result holds for any $\mathbf{a}$. This is the definition of a positive semidefinite matrix, from which it follows that $\boldsymbol{\Sigma}_{\mathbf{x}}$ is positive semidefinite.