Dynamic Systems and Control
Homework 1 Solutions

Ali Vaziri

February 13, 2025

Question 1 Partitioned Matrices

Suppose

$$A = \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix}$$

where the A_i are matrices of conformable dimension.

(a) What can A be premultiplied by to get the matrix

$$\begin{bmatrix} A_3 & A_4 \\ A_1 & A_2 \end{bmatrix}?$$

Solution

$$\begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} A_1 & A_2 \\ A_3 & A_4 \end{bmatrix} = \begin{bmatrix} A_3 & A_4 \\ A_1 & A_2 \end{bmatrix}.$$

(b) Assume that A_1 is nonsingular. What can A be premultiplied by to get the matrix

$$\begin{bmatrix} A_1 & A_2 \\ 0 & C \end{bmatrix}$$

where $C = A_4 - A_3 A_1^{-1} A_2$?

Solution

We use the results from the previous section with the goal of finding a matrix

$$\begin{bmatrix} B_1 & B_2 \\ B_3 & B_4 \end{bmatrix}$$

such that we have

$$BA = \begin{bmatrix} A_1 & A_2 \\ 0 & A_4 - A_3 A_1^{-1} A_2 \end{bmatrix}.$$

This gives us four equations to solve for:

1. $B_1 A_1 + B_2 A_3 = A_1,$
2. $B_1 A_2 + B_2 A_4 = A_2,$
3. $B_3 A_1 + B_4 A_3 = 0,$
4. $B_3 A_2 + B_4 A_4 = A_4 - A_3 A_1^{-1} A_2.$

which can be simply solved to give

$$B = \begin{bmatrix} I & 0 \\ -A_3 A_1^{-1} & I \end{bmatrix}.$$

(c) Suppose A is a square matrix. Use the result in (b) and the fact mentioned in the hint to obtain an expression for $\det(A)$ in terms of determinants involving only the submatrices A_1, A_2, A_3, and A_4.

Solution

Using linear row operations:

$$\det(B) = 1 \quad \text{and} \quad \det(A) = \det(B)\det(A) = \det(A_1)\det(A_4 - A_3 A_1^{-1} A_2).$$

Question 2 Matrix Identities

Prove the following very useful matrix identities. In proving identities such as these, see if you can obtain proofs that make as few assumptions as possible beyond those implied by the problem statement. For example, in (1) and (2) below, neither A nor B need be square, and in (3) neither B nor D need be square — so avoid assuming that any of these matrices is (square and) invertible!

(a) Prove that
$$\det(I - AB) = \det(I - BA),$$
if A is $p \times q$ and B is $q \times p$.

Solution

We begin with the determinant of $I - BA$:

$$\det(I - BA) = \det \begin{pmatrix} I & 0 \\ B & I - BA \end{pmatrix} = \det\{\begin{pmatrix} I & A \\ B & I \end{pmatrix} \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix}\}.$$

$$\det(I - BA) = \det \begin{pmatrix} I & A \\ B & I \end{pmatrix} \cdot \det \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix}.$$

From the previous question, we can write

$$\det \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix} = \det(I) \cdot \det(I) = 1.$$

Thus, we simplify:
$$\det(I - BA) = \det \begin{pmatrix} I & A \\ B & I \end{pmatrix}.$$

Also, we can write

$$\det \begin{pmatrix} I & A \\ B & I \end{pmatrix} = \det \begin{pmatrix} I - AB & 0 \\ B & I \end{pmatrix} = \det(I - AB).$$

Therefore:
$$\det(I - BA) = \det(I - AB).$$

Note: This will be very useful for determinant calculation when we are facing large values for p since $(I - BA)$ is a $q \times q$ matrix, while $(I - AB)$ is a $p \times p$ matrix. To compute the determinant of $(I - AB)$ or $(I - BA)$, compare p and q and choose the product (AB or BA) with the smaller size.

(b) Show that
$$(I - AB)^{-1}A = A(I - BA)^{-1}.$$

Solution

Assume that $(I - BA)$ and $(I - AB)$ are invertible. Then, we start with:

$$A = A \cdot I = A(I - BA)(I - BA)^{-1}.$$

Expanding:
$$A = (A - ABA)(I - BA)^{-1}.$$

Factoring:
$$A = (I - AB)A(I - BA)^{-1}.$$

Thus:
$$(I - AB)^{-1}A = A(I - BA)^{-1}.$$

(c) Show that

$$(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}.$$

Solution

We verify the identity by multiplying both sides by $(A + BCD)$ and showing the result is the identity matrix.

0.1 Left Multiplication

$$\begin{aligned}
&(A + BCD)\left[A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}\right] \\
&= (A + BCD)A^{-1} - (A + BCD)A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}
\end{aligned}$$

0.1.1 First Term Simplification

$$\begin{aligned}(A + BCD)A^{-1} &= AA^{-1} + BCDA^{-1} \\ &= I + BCDA^{-1}\end{aligned}$$

0.1.2 Second Term Simplification

$$\begin{aligned}(A + BCD)A^{-1}B &= [AA^{-1}B + BCDA^{-1}B] \\ &= B + BCDA^{-1}B \\ &= B[I + CDA^{-1}B]\end{aligned}$$

Substituting back:

$$\begin{aligned}B[I + CDA^{-1}B](C^{-1} + DA^{-1}B)^{-1}DA^{-1} &= B[C(C^{-1} + DA^{-1}B)](C^{-1} + DA^{-1}B)^{-1}DA^{-1} \\ &= BCDA^{-1}\end{aligned}$$

0.1.3 Combining Terms

$$I + BCDA^{-1} - BCDA^{-1} = I$$

which is equal to the left-hand side of the equation $(A+BCD)(A+BCD)^{-1} = I$.

This completes the proof.

Question 3 Vandermonde Matrix

We aim to prove that the determinant of the $n \times n$ Vandermonde matrix

$$V = \begin{pmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{pmatrix}$$

is given by

$$\det(V) = \prod_{1 \leq i < j \leq n} (x_j - x_i).$$

We proceed with giving two proofs.

Proof 1. We start with

$$V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-2} & x_2^{n-1} \\ 1 & x_3 & x_3^2 & \cdots & x_3^{n-2} & x_3^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-2} & x_n^{n-1} \end{vmatrix}.$$

By the property that adding a multiple of one row to another does not change the determinant, subtract the first row from each of the other rows:

$$V_n = \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-2} & x_1^{n-1} \\ 0 & x_2 - x_1 & x_2^2 - x_1^2 & \cdots & x_2^{n-2} - x_1^{n-2} & x_2^{n-1} - x_1^{n-1} \\ 0 & x_3 - x_1 & x_3^2 - x_1^2 & \cdots & x_3^{n-2} - x_1^{n-2} & x_3^{n-1} - x_1^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_n - x_1 & x_n^2 - x_1^2 & \cdots & x_n^{n-2} - x_1^{n-2} & x_n^{n-1} - x_1^{n-1} \end{vmatrix}.$$

Next, without changing the determinant's value, subtract in sequence:

- x_1 times the $(n-1)$th column from the nth column,
- x_1 times the $(n-2)$th column from the $(n-1)$th column,
- $\vdots$

- x_1 times the first column from the second column.

After these column operations the first row becomes

$$(1,\ 0,\ 0,\ \ldots,\ 0),$$

and for each row $i \geq 2$ the new entry in column j (for $j \geq 2$) is

$$a_{ij} = \left(x_i^{j-1} - x_1^{j-1}\right) - \left(x_1\, x_i^{j-2} - x_1^{j-1}\right) = (x_i - x_1)x_i^{j-2}.$$

Thus,

$$V_n = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 & 0 \\ 0 & x_2 - x_1 & (x_2 - x_1)x_2 & \cdots & (x_2 - x_1)x_2^{n-3} & (x_2 - x_1)x_2^{n-2} \\ 0 & x_3 - x_1 & (x_3 - x_1)x_3 & \cdots & (x_3 - x_1)x_3^{n-3} & (x_3 - x_1)x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & x_n - x_1 & (x_n - x_1)x_n & \cdots & (x_n - x_1)x_n^{n-3} & (x_n - x_1)x_n^{n-2} \end{vmatrix}.$$

Extract the common factor $(x_i - x_1)$ from each row $i = 2, \ldots, n$:

$$V_n = \left(\prod_{k=2}^{n}(x_k - x_1)\right) \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & x_2 & \cdots & x_2^{n-2} \\ 0 & 1 & x_3 & \cdots & x_3^{n-2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 1 & x_n & \cdots & x_n^{n-2} \end{vmatrix}.$$

Notice that the remaining determinant is precisely the Vandermonde determinant of order $n-1$, denoted by V_{n-1}. Hence,

$$V_n = \left(\prod_{k=2}^{n}(x_k - x_1)\right) V_{n-1}.$$

Since one can directly compute

$$V_2 = \begin{vmatrix} 1 & x_1 \\ 1 & x_2 \end{vmatrix} = x_2 - x_1,$$

the result follows.

Proof 2. We now present an induction proof using an alternative formulation. Write the Vandermonde determinant as

$$V_n = \begin{vmatrix} x_1^{n-1} & x_1^{n-2} & \cdots & x_1 & 1 \\ x_2^{n-1} & x_2^{n-2} & \cdots & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_n^{n-1} & x_n^{n-2} & \cdots & x_n & 1 \end{vmatrix}.$$

Define the proposition $\mathcal{P}(n)$ by

$$\mathcal{P}(n): \quad V_n = \prod_{1 \le i < j \le n} (x_i - x_j).$$

Basis: For $n = 1$, we have $V_1 = 1$. For $n = 2$,

$$V_2 = \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} = x_1 - x_2.$$

Thus, the base case holds.

Induction Hypothesis: Assume that for some $k \ge 2$,

$$V_k = \prod_{1 \le i < j \le k} (x_i - x_j).$$

Induction Step: We must show

$$V_{k+1} = \prod_{1 \le i < j \le k+1} (x_i - x_j).$$

Consider the determinant

$$V'_{k+1} = \begin{vmatrix} x^k & x^{k-1} & \cdots & x^2 & x & 1 \\ x_2^k & x_2^{k-1} & \cdots & x_2^2 & x_2 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ x_{k+1}^k & x_{k+1}^{k-1} & \cdots & x_{k+1}^2 & x_{k+1} & 1 \end{vmatrix},$$

where the variable x replaces x_1. By the Expansion Theorem for Determinants, when we expand along the first row, the determinant becomes a polynomial $f(x)$ in x of degree at most k.

If we substitute $x = x_r$ (with $r = 2, 3, \ldots, k+1$), then the first row equals the rth row, and the determinant vanishes:

$$f(x_r) = 0 \quad \text{for } r = 2, \ldots, k+1.$$

Thus, $f(x)$ has the factors $x - x_2, x - x_3, \ldots, x - x_{k+1}$. In other words,

$$f(x) = C \prod_{r=2}^{k+1} (x - x_r),$$

where C is independent of x. The leading coefficient C is given by the coefficient of x^k; it equals the determinant

$$\begin{vmatrix} x_2^{k-1} & x_2^{k-2} & \cdots & x_2 & 1 \\ x_3^{k-1} & x_3^{k-2} & \cdots & x_3 & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ x_{k+1}^{k-1} & x_{k+1}^{k-2} & \cdots & x_{k+1} & 1 \end{vmatrix}.$$

By the induction hypothesis, this determinant equals

$$\prod_{2 \le i < j \le k+1} (x_i - x_j).$$

Hence, $C = \prod_{2 \le i < j \le k+1} (x_i - x_j)$. Now, substitute $x = x_1$ into $f(x)$ to obtain

$$V_{k+1} = f(x_1) = \left(\prod_{r=2}^{k+1} (x_1 - x_r) \right) \prod_{2 \le i < j \le k+1} (x_i - x_j).$$

Rearranging the factors shows that

$$V_{k+1} = \prod_{1 \le i < j \le k+1} (x_i - x_j).$$

This completes the induction step and hence the proof. ■

Question 4 Matrix Derivatives

(a)

Question: Suppose $A(t)$ and $B(t)$ are matrices whose entries are differentiable functions of t, and assume the product $A(t)B(t)$ is well-defined. Show that

$$\frac{d}{dt}\big(A(t)B(t)\big) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt},$$

where the derivative of a matrix is, by definition, the matrix of derivatives — i.e., to obtain the derivative of a matrix, simply replace each entry of the matrix by its derivative. (Note: The ordering of the matrices in the above result is important!)

Solution: We use the definition of the element-wise derivative in terms of limits:

$$\frac{d}{dt}\big(A(t)B(t)\big) = \lim_{\Delta t \to 0} \frac{A(t+\Delta t)B(t+\Delta t) - A(t)B(t)}{\Delta t}.$$

Note that we can write (using the Taylor series)

$$A(t+\Delta t) = A(t) + \Delta t \frac{dA(t)}{dt} + o(\Delta t),$$
$$B(t+\Delta t) = B(t) + \Delta t \frac{dB(t)}{dt} + o(\Delta t),$$

Now substitute the first-order Taylor series to obtain

$$\begin{aligned}\frac{d}{dt}\big(A(t)B(t)\big) = \lim_{\Delta t \to 0} \frac{1}{\Delta t}\big[A(t)B(t) + \Delta t \frac{dA(t)}{dt}B(t) + \Delta t A(t)\frac{dB(t)}{dt} \\ + \big[A(t) + \Delta t \frac{dA(t)}{dt} + B(t) + \Delta t \frac{dB(t)}{dt}\big] o(\Delta t) + o(\Delta t^2) - A(t)B(t)\big].\end{aligned} \tag{1}$$

Here, $o(\Delta t)$ and $o(\Delta t^2)$ are terms such that

$$\lim_{\Delta t \to 0} \frac{o(\Delta t)}{\Delta t} = 0 \quad \text{and} \quad \lim_{\Delta t \to 0} \frac{o(\Delta t^2)}{\Delta t^2} = 0.$$

For the above expression, we can write it as

$$\underbrace{\big[A(t) + \Delta t\, A'(t)\big]\, o(\Delta t)}_{\text{Term I}} + \underbrace{o(\Delta t)\, \big[B(t) + \Delta t\, B'(t)\big]}_{\text{Term II}} + \underbrace{o(\Delta t^2)}_{\text{Term III}}.$$

We will analyze each term individually and show that the entire expression satisfies

$$\lim_{\Delta t \to 0} \frac{\big[A(t) + \Delta t \frac{dA(t)}{dt} + B(t) + \Delta t \frac{dB(t)}{dt}\big] o(\Delta t) + o(\Delta t^2)}{\Delta t} = 0.$$

Term I: $\big[A(t) + \Delta t\, A'(t)\big]\, o(\Delta t)$

Expand the first term

$$\text{Term I} = \big[A(t) + \Delta t\, A'(t)\big]\, o(\Delta t).$$

As $\Delta t \to 0$:

- The matrix $\big[A(t) + \Delta t\, A'(t)\big]$ is bounded (since $A(t)$ and $A'(t)$ are continuous).
- The product with $o(\Delta t)$ yields $o(\Delta t)$ because multiplying a bounded term by $o(\Delta t)$ does not change the asymptotic order.

Thus:

$$\frac{\text{Term I}}{\Delta t} = \frac{o(\Delta t)}{\Delta t} \to 0 \quad \text{as } \Delta t \to 0.$$

Term II: $o(\Delta t)\,\big[B(t) + \Delta t\, B'(t)\big]$

Expand the second term

$$\text{Term II} = o(\Delta t)\,\big[B(t) + \Delta t\, B'(t)\big].$$

As $\Delta t \to 0$:

- The matrix $\big[B(t) + \Delta t\, B'(t)\big]$ is bounded (since $B(t)$ and $B'(t)$ are continuous).
- The product with $o(\Delta t)$ yields $o(\Delta t)$.

Thus:

$$\frac{\text{Term II}}{\Delta t} = \frac{o(\Delta t)}{\Delta t} \to 0 \quad \text{as } \Delta t \to 0.$$

Term III: $o(\Delta t^2)$

The third term is

$$\text{Term III} = o(\Delta t^2).$$

By definition, $\frac{o(\Delta t^2)}{\Delta t^2} \to 0$ as $\Delta t \to 0$. Hence

$$\frac{\text{Term III}}{\Delta t} = \frac{o(\Delta t^2)}{\Delta t^2} \cdot \Delta t \to 0 \quad \text{as } \Delta t \to 0.$$

Combining all three terms, we have

$$\frac{\text{h.o.t.}}{\Delta t} = \frac{\text{Term I}}{\Delta t} + \frac{\text{Term II}}{\Delta t} + \frac{\text{Term III}}{\Delta t}.$$

Each term goes to 0 as $\Delta t \to 0$. Therefore

$$\lim_{\Delta t \to 0} \frac{\big[A(t) + \Delta t \frac{dA(t)}{dt} + B(t) + \Delta t \frac{dB(t)}{dt}\big] o(\Delta t) + o(\Delta t^2)}{\Delta t} = 0.$$

By reducing the expression in (1) and taking the limit, we obtain

$$\frac{d}{dt}\big(A(t)B(t)\big) = \frac{dA(t)}{dt}B(t) + A(t)\frac{dB(t)}{dt}.$$

which completes the proof.

(b)

Question: Use the result of (a) to evaluate the derivative of the inverse of a matrix $A(t)$, i.e., evaluate the derivative of $A^{-1}(t)$.

Question: We use the identity

$$A^{-1}(t)A(t) = I.$$

Taking the derivative on both sides, we have

$$\frac{d}{dt}\big[A^{-1}(t)A(t)\big] = \frac{d}{dt}A^{-1}(t)A(t) + A^{-1}(t)\frac{dA(t)}{dt} = 0.$$

Multiplying from the right by $A^{-1}(t)$, we obtain

$$\frac{d}{dt}A^{-1}(t) = -A^{-1}(t)\frac{dA(t)}{dt}A^{-1}(t).$$

Question 5 Infinity Norm

Question. Prove that the function

$$\|x\|_\infty \;=\; \max_{1\le i\le n}\bigl|x_i\bigr|$$

defines a norm on $\mathbb{R}^{n\times 1}$. In the literature, this is called the "infinity norm" or ∞-norm.

Solution

We check the three norm axioms

1. Positivity

Let $x = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$.

- Since each $|x_i| \ge 0$, it follows that $\|x\|_\infty = \max_i |x_i| \ge 0$.
- $\|x\|_\infty = 0$ if and only if $\max_i |x_i| = 0$. This forces $|x_i| = 0$ for all i, so $x = 0$.

2. Absolute Homogeneity

For any scalar $\alpha \in \mathbb{R}$,

$$\|\alpha x\|_\infty \;=\; \max_{1\le i\le n}\bigl|\alpha\, x_i\bigr| \;=\; |\alpha|\;\max_{1\le i\le n}|x_i| \;=\; |\alpha|\;\|x\|_\infty.$$

3. Triangle Inequality

For $x, y \in \mathbb{R}^n$,

$$\|x+y\|_\infty \;=\; \max_{1\le i\le n}\bigl|x_i+y_i\bigr| \;\le\; \max_{1\le i\le n}\bigl(|x_i|+|y_i|\bigr) \;\le\; \max_{1\le i\le n}|x_i| + \max_{1\le i\le n}|y_i| \;=\; \|x\|_\infty+\|y\|_\infty.$$

Thus all norm axioms are satisfied, and $\|\cdot\|_\infty$ is a valid norm on $\mathbb{R}^n$.

Question 6 Cauchy-Schwarz Inequality

Let $\langle \cdot, \cdot \rangle$ be an inner product on a real vector space V. Then, for all $x, y \in V$,

$$\big|\langle x, y\rangle\big| \ \leq\ \sqrt{\langle x, x\rangle}\sqrt{\langle y, y\rangle}.$$

Equivalently,

$$\langle x, y\rangle^2 \ \leq\ \langle x, x\rangle\langle y, y\rangle.$$

Solution

We prove the Cauchy–Schwarz Inequality in a real inner product space. The proof uses the fact that the inner product of a vector with itself is always nonnegative.

For any real scalar α, define the function:

$$f(\alpha) \ =\ \langle x - \alpha y,\ x - \alpha y\rangle.$$

Since $\langle \cdot, \cdot \rangle$ is an inner product, it follows that:

$$f(\alpha) \ \geq\ 0 \quad \text{for all } \alpha \in \mathbb{R}.$$

Expanding $f(\alpha)$ using the linearity and symmetry properties of the inner product:

$$f(\alpha) \ =\ \langle x, x\rangle - 2\alpha\langle x, y\rangle + \alpha^2\langle y, y\rangle.$$

This is a quadratic polynomial in α:

$$f(\alpha) \ =\ \langle y, y\rangle\alpha^2 - 2\langle x, y\rangle\alpha + \langle x, x\rangle.$$

Since $f(\alpha) \geq 0$ for all α, the quadratic equation must have a nonpositive discriminant. The discriminant Δ is given by:

$$\Delta \ =\ \big(-2\langle x, y\rangle\big)^2 - 4\langle y, y\rangle\langle x, x\rangle.$$

Simplify:

$$\Delta \ =\ 4\langle x, y\rangle^2 - 4\langle x, x\rangle\langle y, y\rangle.$$

For $f(\alpha) \geq 0$ to hold, we must have:

$$\Delta \ \leq\ 0.$$

The condition $\Delta \leq 0$ simplifies to:

$$\langle x, y \rangle^2 \;\leq\; \langle x, x \rangle \langle y, y \rangle.$$

Taking square roots on both sides gives:

$$\big|\langle x, y \rangle\big| \;\leq\; \sqrt{\langle x, x \rangle}\sqrt{\langle y, y \rangle}.$$

Equality holds if and only if $f(\alpha)$ has exactly one root, which happens when the discriminant $\Delta = 0$. This implies that x and y are linearly dependent, i.e., there exists a scalar β such that $x = \beta y$.

Question 7

Question. Define the following inner product and norm for $\mathbb{R}^{3\times 1}$:

$$\langle x, y\rangle \;=\; x^\top Q\, y, \quad \|x\| \;=\; \sqrt{\langle x, x\rangle},$$

where $x, y \in \mathbb{R}^{3\times 1}$, and Q is a symmetric positive definite 3×3 matrix. Consider the subspace $M \subset \mathbb{R}^{3\times 1}$ spanned by

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad v_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix},$$

and let $y = \begin{pmatrix} 7 \\ 8 \\ 12 \end{pmatrix}$. Use the projection theorem to find

$$\hat{m} \;=\; \arg\min_{m\in M} \|\, y - m\|$$

for two cases of Q:

$$Q_1 \;=\; \begin{pmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 2 & 0.5 \\ 0.2 & 0.5 & 3 \end{pmatrix} \quad \text{and} \quad Q_2 \;=\; \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Solution

We aim to find the Q-orthogonal projection $\hat{m}$ of a vector y onto a subspace $M \subset \mathbb{R}^n$. The inner product is defined as:

$$\langle x, y\rangle_Q = x^\top Q y,$$

where Q is a symmetric positive definite matrix. The subspace M is spanned by the vectors v_1 and v_2, and y is a given vector. The projection theorem and its derivation are detailed below.

Projection Theorem for Q-Inner Products

1. Q-Orthogonality A vector $u \in \mathbb{R}^n$ is said to be Q-orthogonal to a vector $v \in \mathbb{R}^n$ if

$$\langle u, v\rangle_Q = u^\top Q v = 0.$$

2. Statement of the Projection Theorem Given a subspace $M \subset \mathbb{R}^n$, the Q-orthogonal projection $\hat{m}$ of $y \in \mathbb{R}^n$ onto M is characterized by the property:

$$(y - \hat{m})^\top Q w = 0 \quad \text{for all } w \in M.$$

In other words, the residual $y - \hat{m}$ is Q-orthogonal to every vector in M.

—

Derivation of the Projection Formula

1. Subspace Representation Suppose the subspace M is spanned by $v_1, v_2, \ldots, v_k$. Let

$$V = \begin{pmatrix} v_1 & v_2 & \cdots & v_k \end{pmatrix}$$

be the $n \times k$ matrix whose columns are the basis vectors of M. Any vector $m \in M$ can be written as:

$$m = Vc, \quad \text{for some } c \in \mathbb{R}^k.$$

2. Orthogonality Condition The projection theorem requires that $(y - \hat{m})^\top Q w = 0$ for all $w \in M$. Substituting $w = Vc$ (since Vc represents all vectors in M), the orthogonality condition becomes:

$$(y - \hat{m})^\top Q(Vc) = 0 \quad \text{for all } c \in \mathbb{R}^k.$$

3. Expanding the Expression Let $\hat{m} = Va$, where $a \in \mathbb{R}^k$ is the vector of coefficients for $\hat{m}$ in the basis $v_1, \ldots, v_k$. Substituting $\hat{m} = Va$, the orthogonality condition becomes:

$$(y - Va)^\top Q(Vc) = 0.$$

Expanding this:

$$y^\top Q(Vc) - a^\top V^\top Q(Vc) = 0.$$

Factoring out Vc:

$$\left(y^\top QV - a^\top V^\top QV\right)c = 0.$$

4. Solving for a The above equation must hold for all $c \in \mathbb{R}^k$. This implies:

$$y^\top QV = a^\top V^\top QV.$$

Taking the transpose:

$$V^\top Q y = V^\top QVa.$$

Since $V^\top QV$ is invertible (due to the positive definiteness of Q and the linear independence of $v_1, \ldots, v_k$), we solve for a:

$$a = (V^\top QV)^{-1} V^\top Qy.$$

5. Computing $\hat{m}$ The projection $\hat{m}$ is then given by:

$$\hat{m} = Va = V(V^\top QV)^{-1} V^\top Qy.$$

The matrix $P_Q = V(V^\top QV)^{-1}V^\top Q$ is the Q-orthogonal projection operator, which maps any vector $y \in \mathbb{R}^n$ to its Q-orthogonal projection onto M.

Numerical Solution

We now compute $\hat{m}$ for the specific cases provided. Subspace and vector

$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix}, \quad y = \begin{pmatrix} 7 \\ 8 \\ 12 \end{pmatrix}.$$

$$V = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}.$$

Case 1: Q_1
Step 1: Compute $Q_1 y$

$$Q_1 = \begin{pmatrix} 1 & 0.5 & 0.2 \\ 0.5 & 2 & 0.5 \\ 0.2 & 0.5 & 3 \end{pmatrix}.$$

$$Q_1 y = \begin{pmatrix} 13.4 \\ 25.5 \\ 41.4 \end{pmatrix}.$$

Step 2: Compute $V^\top Q_1 y$

$$V^\top Q_1 y = \begin{pmatrix} 188.6 \\ 429.5 \end{pmatrix}.$$

Step 3: Compute $V^\top Q_1 V$

$$V^\top Q_1 V = \begin{pmatrix} 45.2 & 101.6 \\ 101.6 & 233.6 \end{pmatrix}.$$

Step 4: Compute $(V^\top Q_1 V)^{-1}$

$$(V^\top Q_1 V)^{-1} = \frac{1}{236.16}\begin{pmatrix} 233.6 & -101.6 \\ -101.6 & 45.2 \end{pmatrix}.$$

Step 5: Compute $\hat{m}$

$$\hat{m} = V(V^\top Q_1 V)^{-1} V^\top Q_1 y = \begin{pmatrix} 6.04 \\ 8.88 \\ 11.73 \end{pmatrix}.$$

Case 2: $Q_2 = I$
For $Q_2 = I$, the computation reduces to the standard Euclidean projection:

$$V^\top V = \begin{pmatrix} 14 & 32 \\ 32 & 77 \end{pmatrix}, \quad V^\top y = \begin{pmatrix} 59 \\ 140 \end{pmatrix}.$$

$$(V^\top V)^{-1} = \frac{1}{54}\begin{pmatrix} 77 & -32 \\ -32 & 14 \end{pmatrix}.$$

$$\hat{m} = V(V^\top V)^{-1} V^\top y = \begin{pmatrix} 6.5 \\ 9 \\ 11.5 \end{pmatrix}.$$

Final Results

1. For Q_1:

$$\boxed{\hat{m} = \begin{pmatrix} 6.04 \\ 8.88 \\ 11.73 \end{pmatrix}}$$

2. For $Q_2 = I$:

$$\boxed{\hat{m} = \begin{pmatrix} 6.5 \\ 9 \\ 11.5 \end{pmatrix}}$$

Question 8

Question. Define the following inner product and norm on $C[0,1]$, the space of continuous functions on the interval $[0,1]$:

$$\langle f(x), g(x)\rangle \;=\; \int_0^1 f(x)\,g(x)\,dx, \quad \|\,f(x)\| \;=\; \sqrt{\langle f(x), f(x)\rangle}.$$

Now consider the subspace $P \subset C[0,1]$ consisting of all polynomials of degree ≤ 3, spanned by the basis $\{1,\, x,\, x^2,\, x^3\}$. We want to approximate

$$h(x) = e^x$$

by a polynomial $p(x) \in P$. Use the *projection theorem* to find the best polynomial

$$\hat{h}(x) \;=\; \arg\min_{p(x)\in P} \;\|\,h(x) - p(x)\|.$$

Verify your result by comparing the plots of $h(x)$ and $\hat{h}(x)$ on $[0,1]$.

Solution Using Gram-Schmidt Process

Step 1: Construct an Orthogonal Basis for P_3

The standard basis for P_3 is $\{1, x, x^2, x^3\}$. We apply the Gram-Schmidt process to orthogonalize it:

1. **First basis vector:**
$$p_0(x) = 1.$$

2. **Second basis vector:**
$$p_1(x) = x - \frac{\langle x, p_0(x)\rangle}{\langle p_0(x), p_0(x)\rangle} p_0(x) = x - \frac{1/2}{1}\cdot 1 = x - \frac{1}{2}.$$

3. **Third basis vector:**
$$p_2(x) = x^2 - \frac{\langle x^2, p_0(x)\rangle}{\langle p_0(x), p_0(x)\rangle} p_0(x) - \frac{\langle x^2, p_1(x)\rangle}{\langle p_1(x), p_1(x)\rangle} p_1(x).$$

   Calculate:
$$\langle x^2, 1\rangle = \frac{1}{3}, \quad \langle x^2, p_1(x)\rangle = \frac{1}{12}, \quad \langle p_1(x), p_1(x)\rangle = \frac{1}{12}.$$

Thus:

$$p_2(x) = x^2 - \frac{1}{3} - \left(x - \frac{1}{2}\right) = x^2 - x + \frac{1}{6}.$$

4. **Fourth basis vector:**

$$p_3(x) = x^3 - \frac{\langle x^3, p_0(x)\rangle}{\langle p_0(x), p_0(x)\rangle} p_0(x) - \frac{\langle x^3, p_1(x)\rangle}{\langle p_1(x), p_1(x)\rangle} p_1(x) - \frac{\langle x^3, p_2(x)\rangle}{\langle p_2(x), p_2(x)\rangle} p_2(x).$$

After computing inner products and simplifying:

$$p_3(x) = x^3 - \frac{3}{2}x^2 + \frac{3}{5}x - \frac{1}{20}.$$

Step 2: Project $h(x) = e^x$ onto the Orthogonal Basis

The best approximation is the projection of e^x to the orthogonal basis. This gives the closest function to the $h(x)$.

$$\hat{h}(x) = Proj_{\{}p_0, p_1, p_2, p_3\}(e^x) = \sum_{k=0}^{3} \frac{\langle e^x, p_k(x)\rangle}{\langle p_k(x), p_k(x)\rangle} p_k(x).$$

1. **Compute coefficients:**

$$\frac{\langle e^x, p_0(x)\rangle}{\langle p_0(x), p_0(x)\rangle} = e - 1, \quad \frac{\langle e^x, p_1(x)\rangle}{\langle p_1(x), p_1(x)\rangle} = 6(3 - e),$$
$$\frac{\langle e^x, p_2(x)\rangle}{\langle p_2(x), p_2(x)\rangle} = 0.84, \quad \frac{\langle e^x, p_3(x)\rangle}{\langle p_3(x), p_3(x)\rangle} = 042.$$

2. **Combine terms:**

$$\hat{h}(x) = 0.42x^3 + 0.195x^2 + 1.205x + 0.946$$

Final Answer

The best polynomial approximation of $h(x) = e^x$ in P_3 is:

$$\boxed{\hat{h}(x) = 0.42x^3 + 0.195x^2 + 1.205x + 0.946}.$$

Verification

A plot of e^x and $\hat{h}(x)$ over $[0, 1]$ would show close agreement, confirming the accuracy of the approximation.

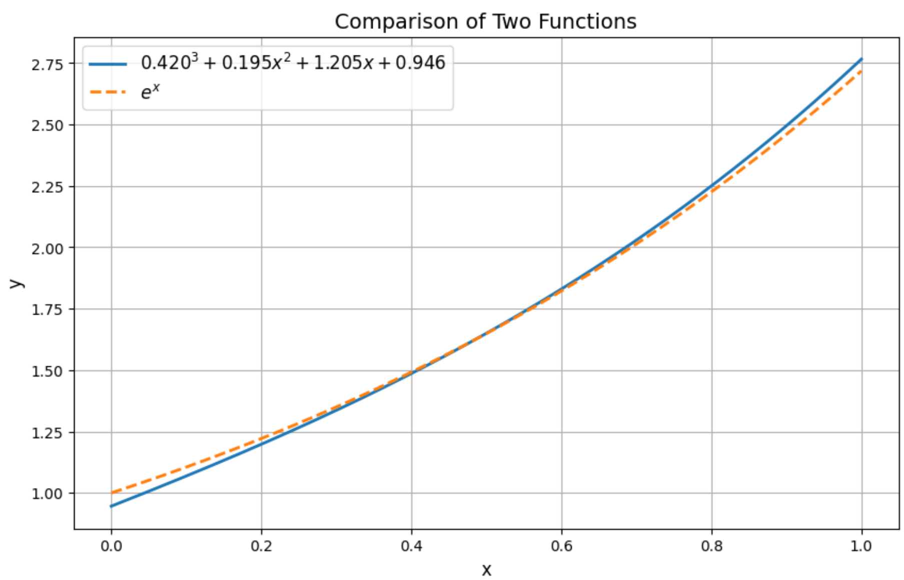

Figure 1: Closeness of the function approximation.