MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Department of Electrical Engineering and Computer Science

6.241: Dynamic Systems—Fall 2003

Homework 1 Solutions

Exercise 1.1 a)Given A_1 and A_4 are square matrices, we know that A is square:

$$A = \begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & A_4 \end{pmatrix}.$$

Note that

$$det \begin{pmatrix} A_1 & A_2 \\ 0 & I \end{pmatrix} = det(I)det(A_4) = det(A_4),$$

which can be verified by recursively computing the principal minors. Also, by the elementary operations of rows, we have

$$det = \begin{pmatrix} A_1 & A_2 \\ 0 & I \end{pmatrix} = det \begin{pmatrix} A_1 & 0 \\ 0 & I \end{pmatrix} = det(A_1).$$

Thus, finally using the fact that $det(AB) = det(A) = det(B)$, we have

$$det(A) = det(A_1)det(A_4).$$

b) Assume A_1^{-1} and A_4^{-1} exist. Then

$$AA^{-1} = \begin{pmatrix} A_1 & A_2 \\ 0 & A_4 \end{pmatrix} \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix} = \begin{pmatrix} I & 0 \\ 0 & I \end{pmatrix},$$

which yields four matrix equations:

1. $A_1B_1 + A_2B_3 = I$,
2. $A_1B_2 + A_2B_4 = 0$,
3. $A_4B_3 = 0$,
4. $A_4B_4 = I$.

From Eqn (4), $B_4 = A_4^{-1}$, with which Eqn (2) yields $B_2 = -A_1^{-1}A_2A_4^{-1}$. Also, from Eqn (3) $B_3 = 0$, with which from Eqn (1) $B_1 = A_1^{-1}$. Therefore,

$$A^{-1} = \begin{pmatrix} A_1^{-1} & -A_1^{-1}A_2A_4^{-1} \\ 0 & A_4^{-1} \end{pmatrix}.$$

Exercise 1.2 a)

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

b) Let us find

$$B = \begin{pmatrix} B_1 & B_2 \\ B_3 & B_4 \end{pmatrix}$$

such that

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

The above equation implies four equations for submatrices

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. $B_3 A_2 + B_4 A_4 = A_4 - A_3 A_1^{-1} A_2$.

First two equations yield $B_1 = I$ and $B_2 = 0$. Express B_3 from the third equation as $B_3 = -B_4 A_3 A_1^{-1}$ and plug it into the fourth. After gathering the terms we get $B_4 \left(A_4 - A_3 A_1^{-1} A_2\right) = A_4 - A_3 A_1^{-1} A_2$, which turns into identity if we set $B_4 = I$. Therefore

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

c) Using linear operations on rows we see that $det\,(B) = 1$. Then, $det(A) = det(B)det(A) = det\,(BA) = det\,(A_1)\,det\left(A_4 - A_3 A_1^{-1} A_2\right)$. Note that $\left(A_4 - A_3 A_1^{-1} A_2\right)$ does not have to be invertible for the proof.

Exercise 1.3 We have to prove that $det(I - AB) = det(I - BA)$.
Proof: Since I and $I - BA$ are square,

$$\begin{aligned} det(I - BA) &= det \begin{pmatrix} I & 0 \\ B & I - BA \end{pmatrix} \\ &= det\left(\begin{pmatrix} I & A \\ B & I \end{pmatrix} \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix} \right) \\ &= det \begin{pmatrix} I & A \\ B & I \end{pmatrix} \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix}, \end{aligned}$$

yet, from Exercise 1.1, we have

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

Thus,

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

Now,

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

Therefore

$$det(I - BA)det(I - AB).$$

Note that $(I - BA)$ is a $q \times q$ matrix while $(I - AB)$ is a $p \times p$ matrix. Thus, one wants to compute the determinant of $(I - AB)$ or $(I - BA)$, s/he compare p and q to pick a smaller size of product either AB or BA.
b) We have to show that $(I - AB)^{-1}A = A(I - BA)^{-1}$.
Proof: Assume that $(I - BA)^{-1}$ and $(I - AB)^{-1}$ exist. Then,

$$\begin{aligned} A &= A \cdot I = A(I - BA)(I - BA)^{-1} \\ &= (A - BAB)(I - BA)^{-1} \\ &= (I - AB)A(I - BA)^{-1} \\ \to (I - AB)^{-1}A &= A(I - BA)^{-1}. \end{aligned}$$

This completes the proof.
c) We have to show that $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$.
Proof: Suppose $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$, then

$$\begin{aligned} (A + BCD)^{-1}(A + BCD) &= I = (A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1})(A + BCD) \\ \\ &= (I + A^{-1}BCD) - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}D(I + A^{-1}BCD) \\ &= (I + A^{-1}BCD) - A^{-1}B(C^{-1}(I + CDA^{-1}B))^{-1}D(I + A^{-1}BCD) \\ &= (I + A^{-1}BCD) - A^{-1}B(I + CDA^{-1}B)^{-1}CD(I + A^{-1}BCD) \\ &= (I + A^{-1}BCD) - A^{-1}BCD(I + A^{-1}BCD)^{-1}(I + A^{-1}BCD) \\ &= I \because \text{ part b.} \end{aligned}$$

where we have used from part (b) that $(I + CDA^{-1}B)^{-1}CD = CD(I + A^{-1}BCD)^{-1}$. Then we have to show that $(I - ab^T)^{-1}$only requires inversion of A scalar.
Proof: From the lemma proved above, let

$$A = I, \quad B = -a, \quad C = I, \quad D = b^T,$$

then

$$\begin{aligned} (I - ab^T)^{-1} &= I + a(I - b^T a)^{-1}b^T \\ &= I + \frac{ab^T}{(1 - b^T a)}. \end{aligned}$$

Note that $b^T a$ is a scalar.

Exercise 1.4 a) First define all the spaces:

$$\begin{aligned} \mathcal{R}(A) &= \{y \in \mathbf{C}^n \mid Ax = y, \forall x \in \mathbf{C}^n\} \\ \mathcal{R}^{\perp}(A) &= \{z \in \mathbf{C}^m \mid y'z = z'y = 0, \forall y \in \mathcal{R}(A)\} \\ \mathcal{R}(A') &= \{p \in \mathbf{C}^n \mid A'v = p, \forall v \in \mathbf{C}^m\} \\ \mathcal{N}(A) &= \{x \in \mathbf{C}^n \mid Ax = 0\} \\ \mathcal{N}(A') &= \{q \in \mathbf{C}^m \mid A'q = 0\} \end{aligned}$$

i) Prove that $\mathcal{R}^{\perp}(A) = \mathcal{N}(A')$.
Proof: Let

$$\begin{aligned} z \in \mathcal{R}^{\perp}(A) &\rightarrow y'z = 0 \ \forall y \in \mathcal{R}(A) \\ &\rightarrow x'A'z = 0 \ \forall x \in \mathbf{C}^n \\ &\rightarrow A'z = 0 \rightarrow z \in \mathcal{N}(A') \\ &\rightarrow \mathcal{R}^{\perp}(A) \subset \mathcal{N}(A'). \end{aligned}$$

Now let

$$\begin{aligned} q \in \mathcal{N}(A') &\rightarrow A'q = 0 \\ &\rightarrow x'A'q = 0 \ \forall x \in \mathbf{C}^n \\ &\rightarrow y'q = 0 \ \forall y \in \mathcal{R}(A) \\ &\rightarrow q \in \mathcal{R}^{\perp}(A) \\ &\rightarrow \mathcal{N}(A') \subset \mathcal{R}^{\perp}(A). \end{aligned}$$

Therefore

$$\mathcal{R}^{\perp}(A) = \mathcal{N}(A').$$

ii) Prove that $\mathcal{N}^{\perp}(A) = \mathcal{R}(A')$.
Proof: From i) we know that $\mathcal{N}(A) = \mathcal{R}^{\perp}(A')$ by switching A with A'. That implies that

$$\mathcal{N}^{\perp}(A) = \{\mathcal{R}^{\perp}(A')\}^{\perp} = \mathcal{R}(A').$$

b) Show that $rank(A) + rank(B) - n \leq rank(AB) \leq \min\{rank(A), rank(B)\}$.
Proof: i) Show that $rank(AB) \leq min\{rank(A), rank(B)\}$. It can be proved as follows:
Each column of AB is a combination of the columns of A, which implies that $\mathcal{R}(AB) \subseteq \mathcal{R}(A) \rightarrow rank(AB) \leq rank(A)$.
Each row of AB is a combination of the rows of B $\rightarrow$ rowspace $(AB) \subseteq$ rowspace (B), but the dimension of rowspace = dimension of column space = rank, so that $rank(AB) \leq rank(B)$.
Therefore,

$$rank(AB) \leq min\{rank(A), rank(B)\}.$$

ii) Show that $rank(A) + rank(B) - n \leq rank(AB)$.
Let

$$\begin{aligned} r_B &= rank(B) \\ r_A &= rank(A) \\ K_B &= dim(\mathcal{N}(B')) = \text{ nullity of } B' = n - r_B \\ K_A &= \text{nullity of } A = n - r_A, \end{aligned}$$

where $A \in \mathbf{C}^{m \times n}$, $B \in \mathbf{B}^{n \times p}$.
Now, let $\{v_1, \cdots, v_{r_B}\}$ be a basis set of $\mathcal{R}(B)$, and add $n - r_B$ linearly independent vectors $\{w_1, \cdots, w_{n-r_B}\}$ to the basis to span all of $\mathbf{C}^n$, $\{v_1, v_2, \cdots, v_n, w_1, \cdots, w_{n-r_B}\}$.
Let

$$M = \begin{pmatrix} v_1| & v_2 & \cdots & v_{r_B}| & w_1| & \cdots & w_{n-r_B} \end{pmatrix} = \begin{pmatrix} V & W \end{pmatrix}.$$

Suppose $x \in \mathbf{C}^n$, then $x = M\alpha$ for some $\alpha \in \mathbf{C}^n$.

1. $\mathcal{R}(A) = \mathcal{R}(AM) = \mathcal{R}([AV|AW])$.
   Proof: i) Let $x \in \mathcal{R}(A) \rightarrow Ay = x$ for some $y \in \mathbf{C}^n$. Yet , y can be written as a linear combination of the basis vectors of $\mathbf{C}^n$, so $y = M\alpha$ for some $\alpha \in \mathbf{C}^n$.
   Then, $Ay = AM\alpha = x \rightarrow x \in \mathcal{R}(AM) \rightarrow \mathcal{R}(A) \subset \mathcal{R}(AM)$.
   ii) Let $x \in \mathcal{R}(AM) \rightarrow AMy = x$ for some $y \in \mathcal{C}^n$, but $My = z \in \mathbf{C}^n \rightarrow Az = x \rightarrow x \in \mathcal{R}(A) \rightarrow \mathcal{R}(AM) \subset \mathcal{R}(A)$.
   Therefore, $\mathcal{R}(A) = \mathcal{R}(AM) = \mathcal{R}([AV|AW])$.

2. $\mathcal{R}(AB) = \mathcal{R}(AV)$.
   Proof: i) Let $x \in \mathcal{R}(AV) \rightarrow AVy = x$ for some $y \in \mathbf{C}^{r_B}$. Yet, $Vy = B\alpha$ for some $\alpha \in \mathbf{C}^p$ since the columns of V and B span the same space. That implies that $AVy = AB\alpha = x \rightarrow x \in \mathcal{R}(AB) \rightarrow \mathcal{R}(AV) \subset \mathcal{R}(AB)$.
   ii) Let $x \in \mathcal{R}(AB) \rightarrow (AB)y = x$ for some $y \in \mathbf{C}^p$. Yet, again $By = V\theta$ for some $\theta \in \mathbf{C}^{r_B} \rightarrow ABy = AV\theta = x \rightarrow x \in \mathcal{R}(AV) \rightarrow \mathcal{R}(AB) \subset \mathcal{R}(AV)$.
   Therefore, $\mathcal{R}(AV) = \mathcal{R}(AB)$.

Using the fact 1, we see that the number of linearly independent columns of A is less than or equal to the number of linearly independent columns of AV + the number of linearly independent columns of AW. , which means that

$$rank(A) \leq rank(AV) + rank(AW).$$

Using the fact 2, we see that

$$rank(AV) = rank(AB) \rightarrow rank(A) \leq rank(AB) + rank(AW),$$

yet, there re only $n - r_B$ columns in AW. Thus,

$$\begin{aligned} &\rightarrow \quad rank(AW) \leq n - r_B \\ &\rightarrow \quad rank(A) - rank(AB) \leq rank(AW) \leq n - r_B \\ &\rightarrow \quad r_A - (n - r_B) \leq r_{AB}. \end{aligned}$$

This completes the proof.

Exercise 1.8 Let $X = \{g(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_M x^M \mid \alpha_i \in \mathbf{C}\}$.
a) We have to show that the set $B = \{1, x, \cdots, x^M\}$ is a basis for X.
Proof :

1. First, let's show that elements in B are linearly independent. It is clear that each element in B can not be written as a linear combination of each other. More formally,

   $$c_1(1) + c_1(x) + \cdots + c_M(x^M) = 0 \leftrightarrow \forall i \; c_i = 0.$$

   Thus, elements of B are linearly independent.

2. Then, let's show that elements in B span the space X. Every polynomial of order less than or equal to M look like

   $$p(x) = \sum_{i=0}^{M} \alpha_i x^i$$

   for some set of α_i's.
   Therefore, $\{1, x_1, \cdots, x^M\}$ span X.

b) $T : X \to X$ and $T(g(x)) = \frac{d}{dx} g(x)$.

1. Show that T is linear.
   Proof:

$$\begin{aligned} T(ag_1(x) + bg_2(x)) &= \frac{d}{dx}(ag_1(x) + bg_2(x)) \\ &= a\frac{d}{dx}g_1 + b\frac{d}{dx}g_2 \\ &= aT(g_1) + bT(g_2). \end{aligned}$$

   Thus, T is linear.

2. $g(x) = \alpha_0 + \alpha_1 x + \alpha_2 x^2 + \cdots + \alpha_M x^M$, so

$$T(g(x)) = \alpha_1 + 2\alpha_2 x + \cdots + M\alpha_M x^{M-1}.$$

   Thus it can be written as follows:

$$\begin{pmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & 0 & \cdots & M \\ 0 & 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \vdots \\ \alpha_M \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ 3\alpha_3 \\ \vdots \\ M\alpha_M \\ 0 \end{pmatrix}.$$

   The big matrix, M, is a matrix representation of T with respect to basis B. The column vector in the left is a representation of $g(x)$ with respect to B. The column vector in the right is $T(g)$ with respect to basis B.

3. Since the matrix M is upper triangular with zeros along diagonal (in fact M is Hessenberg), the eigenvalues are all 0;

$$\lambda_i = 0 \ \forall i = 1, \cdots, M+1.$$

4. One eigenvector of M for $\lambda_1 = 0$ must satisfy $MV_1 = \lambda_1 V_1 = 0$

$$V_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

   is one eigenvector. Since λ_i 's are not distinct, the eigenvectors are not necessarily independent. Thus in order to computer the M others, ones uses the generalized eigenvector formula.