EE5120 Linear Algebra: Tutorial 4, July-Dec 2017-18

1. State True or False for each of the following with proper justification:
   (a) Let matrix A be a transformation from $\mathbb{R}^m$ to $\mathbb{R}^n$, then dimension of left nullspace of A, i.e. $N(A^T)$ is $m-r$.
   (b) The pseudoinverse $(A^tA)^{-1}A$ of any linear operator A exists even if the operator is not invertible.
   (c) Let V and W be vector spaces, and let $T : V \to W$ be linear. The T is one-to-one iff $N(T) = \{0\}$.
   (d) Let $v \in \mathbb{R}^n$. The nullity of matrix vv^t is n.
   (e) $A = \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$ is a rotation matrix and $B = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ is a reflection matrix.

Solution:
(a) False. Dimension will be $n-r$.

(b) False. A^tA needs to be invertible.

(c) True. Suppose that $\mathbf{T}$ is one-to-one and $x \in N(\mathbf{T})$. Then $\mathbf{T}(x) = 0 = \mathbf{T}(0)$. Hence, $N(T) = \{0\}$.

(d) False. vv^t is a rank 1 matrix. By rank-nullity theorem, nullity $= n-1$.

(e) True.

2. In $\mathbb{R}^2$, let L be the line $y = 2x$. Find an expression for $\mathbf{T}(x, y)$, where $\mathbf{T}$ is the reflection of $\mathbb{R}^2$ about L.

Solution: Refer section 2.6 of Gilbert Strang. The matrix for reflection about θ line is $H = \begin{bmatrix} 2c^2-1 & 2cs \\ 2cs & 2s^2-1 \end{bmatrix}$, where $c = cos\theta$ and $s = sin\theta$. For line $y = 2x$, $\theta = tan^{-1}(2)$.
Substitute these values in H and you get, $\mathbf{T}(x, y)$ = (1/5)$\begin{bmatrix} -3x+4y \\ 4x+3y \end{bmatrix}$.

3. Prove that for two matrices, A, B, the following holds: rank(AB) $\leq$ min(rank(A),rank(B)).

Solution: **Step 1 (col picture)**: Consider the product AB in the following way: $AB = A[b_1 \ldots b_n] = [Ab_1 \ldots Ab_n]$, where b_i is a column of B. Each of these Ab_i is a linear combination of the columns of A, hence, $Ab_i \in C(A)$ (col space of A), thus $C(AB) \in C(A)$. This implies that the rank of AB can not exceed that of A.
Step 2 (row picture): In the product AB, every row is a linear combination* of the rows of B. We know that linear combinations of rows don't change the rank of a matrix, thus, the rank of AB can not exceed that of B. Putting the two steps together, we get the desired result.
*: To see this, $(AB)_{ij} = \sum_k a_{ik}b_{kj}$. So the i^{th} row of (AB) is:
$[\sum_k a_{ik}b_{k1} \quad \sum_k a_{ik}b_{k2} \quad \ldots \quad \sum_k a_{ik}b_{kn}] = \sum_k a_{ik}[b_{k1} \quad b_{k2} \quad \ldots \quad b_{kn}] = \sum_k a_{ik}[b_k]$ where $[b_k]$ is the k^{th} row of B, i.e. a linear combination of the rows.

4. Let T be a linear transformation from R^3 into R^2 and U be a linear transformation from R^2 into R^3. Prove that the transformation UT is not invertible. Generalize the theorem. (Can you relate this to question no.7 of the previous tutorial?)

Solution: Since U, T are linear transformations, they must have matrix representations. T is 2×3, and U is 2×3, thus both their ranks can be at most 2. Using the result above, even though the size of UT is 3×3, it can have rank at most 2. Thus, UT is not invertible.
Alt proof: For a transformation to be invertible, it should be both one-to-one and onto. In this case, T is not one-to-one and therefore UT is not one-to-one. Hence it is not invertible. Any transformation from a higher dimensional space to a lower dimensional space leads to loss of information in one or more dimensions and hence is not invertible.

5. What 3 by 3 matrices represent the transformations that,
(a) project every vector onto the x-y plane?
(b) reflect every vector through the x-y pane?
(c) rotate the x-y plane through 90^o, leaving the z-axis alone?
(d) rotate the x-y plane, then x-z, then y-z through 90^o?
(e) carry out the same three rotations, but each one through 180^o?

Solution:

(a) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

(b) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$

(c) $\begin{bmatrix} cos(\pi/2) & -sin(\pi/2) & 0 \\ sin(\pi/2) & cos(\pi/2) & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

(d) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$

(e) $\begin{bmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$

6. Let $\alpha_1 = \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \right\}$ be the (ordered) basis for the vector space $M_{2\times 2}$, which is the set of all real valued 2×2 matrices. Also, let $\alpha_2 = \{x^2, x, 1\}$ be the basis for the vector space P_2, which is the set of all real polynomials (with real co-efficients) with minimum degree 2. Compute the matrix representations for the following linear transformations:

(a) $\mathsf{T}_1 : \mathsf{M}_{2\times 2} \to \mathsf{M}_{2\times 2}$ with $\mathsf{T}_1(\mathbf{A}) = \mathbf{A}^T$, for every $\mathbf{A} \in \mathsf{M}_{2\times 2}$.

(b) $\mathsf{T}_2 : \mathsf{P}_2 \to \mathsf{M}_{2\times 2}$ with $\mathsf{T}_2(f(x)) = \begin{bmatrix} f'(0) & 2f(1) \\ 0 & f''(3) \end{bmatrix}$. Here, $f'(x)$ and $f''(x)$ are the 1^{st} and 2^{nd} derivatives of $f(x) \in \mathsf{P}_2$.

Solution:

(a) We have the following:

$$\mathsf{T}_1\left(\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix}\right) = \begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} = (1)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (0)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}.$$

We take the co-efficients present in the linear combination shown above to construct the first column in the matrix representation of T_1 will be $[1\,0\,0\,0]^T$. Similarly,

$$\mathsf{T}_1\left(\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\right) = \begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} = (0)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (1)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (0)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix},$$
$$\mathsf{T}_1\left(\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix}\right) = \begin{bmatrix}0 & 0\\1 & 0\end{bmatrix} = (0)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (0)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (1)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}$$
$$\mathsf{T}_1\left(\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}\right) = \begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} = (0)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (1)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}.$$

Thus, matrix representation of T_1 is given by $\begin{bmatrix}1 & 0 & 0 & 0\\0 & 1 & 0 & 0\\0 & 0 & 0 & 1\\0 & 0 & 1 & 0\end{bmatrix}$.

(b) Following the same procedure as discussed above, we get,

$$\mathsf{T}_2(x^2) = \begin{bmatrix}2(0) & 2(1^2)\\0 & 2\end{bmatrix} = \begin{bmatrix}0 & 2\\0 & 2\end{bmatrix} = (0)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (2)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (2)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix},$$
$$\mathsf{T}_2(x) = \begin{bmatrix}1 & 2(1)\\0 & 0\end{bmatrix} = \begin{bmatrix}1 & 2\\0 & 0\end{bmatrix} = (1)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (2)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix},$$
$$\mathsf{T}_2(1) = \begin{bmatrix}0 & 2(1)\\0 & 0\end{bmatrix} = \begin{bmatrix}0 & 2\\0 & 0\end{bmatrix} = (0)\begin{bmatrix}1 & 0\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix} + (2)\begin{bmatrix}0 & 1\\0 & 0\end{bmatrix} + (0)\begin{bmatrix}0 & 0\\1 & 0\end{bmatrix}.$$

The matrix representation for T_2 is given by $\begin{bmatrix}0 & 1 & 0\\2 & 0 & 0\\2 & 2 & 2\\0 & 0 & 0\end{bmatrix}$.

7. Let V be a vector space and $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be a linear transformation. Suppose $\mathbf{x} \in \mathsf{V}$ is such that $\mathsf{T}^k(\mathbf{x}) = \mathbf{0}$, $\mathsf{T}^m(\mathbf{x}) \neq \mathbf{0}$, $\forall \quad 1 \leq m < k$ and $k > 1$, then prove that the set of vectors $\{\mathbf{x}, \mathsf{T}(\mathbf{x}), \mathsf{T}^2(\mathbf{x}), ..., \mathsf{T}^{k-1}(\mathbf{x})\}$ is linearly independent.

Solution: Given $k > 1$. Thus, $\mathsf{T}(\mathbf{x}) \neq \mathbf{0} \Rightarrow \mathbf{x} \neq \mathbf{0}$. Since $\mathsf{T}^k(\mathbf{x}) = \mathbf{0}$, for all $p \geq 1$,

$$\mathsf{T}^{k+p}(\mathbf{x}) = \mathsf{T}^p\left(\mathsf{T}^k(\mathbf{x})\right) = \mathsf{T}^p(\mathbf{0}) = \mathbf{0}. \tag{1}$$

Assume that $\{\mathbf{x}, \mathsf{T}(\mathbf{x}), \mathsf{T}^2(\mathbf{x}), ..., \mathsf{T}^{k-1}(\mathbf{x})\}$ is linearly dependent. Then,

$$a_1\mathbf{x} + a_2\mathsf{T}(\mathbf{x}) + ... + a_k\mathsf{T}^{k-1}(\mathbf{x})\} = 0,$$

with not all a_is being zero, i.e., some a_is are not equal to zero. Now, consider the following:

$$\begin{aligned}
&\mathsf{T}^{k-1}\left(a_1\mathbf{x} + a_2\mathsf{T}(\mathbf{x}) + ... + a_k\mathsf{T}^{k-1}(\mathbf{x})\}\right) = \mathsf{T}^{k-1}(\mathbf{0}) \\
\Rightarrow &a_1\mathsf{T}^{k-1}(\mathbf{x}) + a_2\mathsf{T}^k(\mathbf{x}) + a_3\mathsf{T}^{k+1}(\mathbf{x}) + ... + a_k\mathsf{T}^{2(k-1)}(\mathbf{x}) = \mathbf{0} \\
\Rightarrow &a_1\mathsf{T}^{k-1}(\mathbf{x}) + \mathbf{0} + \mathbf{0} + ... + \mathbf{0} = \mathbf{0}.
\end{aligned}$$

The above result is a consequence of equation (1) and other given information. Since $\mathsf{T}^{k-1}(\mathbf{x}) \neq \mathbf{0}$, $a_1 = 0$. Now,

$$\begin{aligned}
&\mathsf{T}^{k-2}\left(a_1\mathbf{x} + a_2\mathsf{T}(\mathbf{x}) + ... + a_k\mathsf{T}^{k-1}(\mathbf{x})\}\right) = \mathsf{T}^{k-2}(\mathbf{0}) \\
\Rightarrow &a_1\mathsf{T}^{k-2}(\mathbf{x}) + a_2\mathsf{T}^{k-1}(\mathbf{x}) + a_3\mathsf{T}^{k}(\mathbf{x}) + ... + a_k\mathsf{T}^{2k-3}(\mathbf{x}) = \mathbf{0} \\
\Rightarrow &\mathbf{0} + a_2\mathsf{T}^{k-1}(\mathbf{x}) + \mathbf{0} + ... + \mathbf{0} = \mathbf{0}.
\end{aligned}$$

Again, since $\mathsf{T}^{k-1}(\mathbf{x}) \neq \mathbf{0}$, we get $a_2 = 0$. On repeating this procedure, we get $a_i = 0, \forall i = 1, 2, ..., k$, which is contradicting to the initial assumption. Hence, the initial assumption of the set $\{\mathbf{x}, \mathsf{T}(\mathbf{x}), \mathsf{T}^2(\mathbf{x}), ..., \mathsf{T}^{k-1}(\mathbf{x})\}$ being linearly dependent is incorrect. Thus, the above set is linearly independent.

8. BONUS question: **Definition:** Let V be a vector space and $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be a linear transformation on V. A subspace $\mathsf{W} \subset \mathsf{V}$ is said to be T-invariant if for every $\mathbf{w} \in \mathsf{W}$, $\mathsf{T}(\mathbf{w}) \in \mathsf{W}$. Further, if W is T-invariant, define *restriction of T on W* as, $\mathsf{T}_\mathsf{W} : \mathsf{W} \to \mathsf{W}$ such that $\mathsf{T}_\mathsf{W}(\mathbf{w}) = \mathsf{T}(\mathbf{w}), \forall \mathbf{w} \in \mathsf{W}$. Then, prove the following results:

   (a) Subspaces $\{\mathbf{0}\}$, V, $\mathsf{N}(\mathsf{T})$ and $\mathsf{R}(\mathsf{T})$ are T-invariant. Here, $\mathsf{N}(\mathsf{T}) = \{\mathbf{v} \in \mathsf{V} | \mathsf{T}(\mathbf{v}) = \mathbf{0}\}$, and $\mathsf{R}(\mathsf{T}) = \{\mathbf{u} \in \mathsf{V} | \exists \mathbf{x_u} \in \mathsf{V} \text{ s.t } \mathsf{T}(\mathbf{x_u}) = \mathbf{u}\}$ (*The choice of* $\mathbf{x_u}$ *depends on* $\mathbf{u}$).

   (b) For a T-invariant subspace W, the transformation T_W is linear, and $\mathsf{N}(\mathsf{T}_\mathsf{W}) = \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$.

Solution:

(a) (i) Let $\mathsf{U} = \{\mathbf{0}\}$. Its a singleton set. Since T is linear $\mathsf{T}(\mathbf{0}) = \mathbf{0} \Rightarrow \mathsf{T}(\mathbf{0}) \in \mathsf{U}$. Hence, $\mathsf{U} = \{\mathbf{0}\}$ is T-invariant.

(ii) Since T is defined from V to V, for every $\mathbf{v} \in \mathsf{V}$, $\mathsf{T}(\mathbf{v}) \in \mathsf{V} \Rightarrow \mathsf{V}$ is T-invariant.

(iii) Since T is linear $\mathsf{T}(\mathbf{0}) = \mathbf{0}$. Thus, $\mathbf{0} \in \mathsf{N}(\mathsf{T})$ by the definition of $\mathsf{N}(\mathsf{T})$. Further, $\mathsf{T}(\mathbf{v}) = \mathbf{0} \in \mathsf{N}(\mathsf{T})$, for all $\mathbf{v} \in \mathsf{N}(\mathsf{T})$. Hence, $\mathsf{N}(\mathsf{T})$ is T-invariant.

(iv) Let $\mathbf{u} \in \mathsf{R}(\mathsf{T})$. Since, $\mathsf{R}(\mathsf{T}) \subset \mathsf{V}$ (by definition), $\mathbf{u} \in \mathsf{V}$. Thus, $\mathsf{T}(\mathbf{u}) \in \mathsf{R}(\mathsf{T}) \Rightarrow \mathsf{R}(\mathsf{T})$ is T-invariant.

(b) (i) Let $\mathbf{x}, \mathbf{y} \in \mathsf{W}$. Then, $c_1\mathbf{x} + c_2\mathbf{y} \in \mathsf{W}$, for some scalars c_1, c_2 as W is a subspace. W is T-invariant $\Rightarrow \mathsf{T}(\mathbf{x}) \in \mathsf{W}, \mathsf{T}(\mathbf{y}) \in \mathsf{W}, \mathsf{T}(c_1\mathbf{x} + c_2\mathbf{y}) \in \mathsf{W}$. Thus, $\mathsf{T}_\mathsf{W}(\mathbf{x}) = \mathsf{T}(\mathbf{x}), \mathsf{T}_\mathsf{W}(\mathbf{y}) = \mathsf{T}(\mathbf{y})$ and $\mathsf{T}_\mathsf{W}(c_1\mathbf{x} + c_2\mathbf{y}) = \mathsf{T}(c_1\mathbf{x} + c_2\mathbf{y})$. Now, we have the following:

$$\mathsf{T}_\mathsf{W}(c_1\mathbf{x} + c_2\mathbf{y}) = \mathsf{T}(c_1\mathbf{x} + c_2\mathbf{y}) = c_1\mathsf{T}(\mathbf{x}) + c_2\mathsf{T}(\mathbf{y}) = c_1\mathsf{T}_\mathsf{W}(\mathbf{x}) + c_2\mathsf{T}_\mathsf{W}(\mathbf{y}).$$

The above is true for any scalars c_1, c_2 and for any $\mathbf{x}, \mathbf{y} \in \mathsf{W}$. Thus, T_W is linear.

(ii) Let $\mathbf{w} \in \mathsf{N}(\mathsf{T}_\mathsf{W})$. Since $\mathsf{N}(\mathsf{T}_\mathsf{W}) \subset \mathsf{W}$, $\mathbf{w} \in \mathsf{W}$. Further, $\mathsf{T}_\mathsf{W}(\mathbf{w}) = \mathbf{0}$. But by definition, $\mathsf{T}_\mathsf{W}(\mathbf{w}) = \mathsf{T}(\mathbf{w}) \Rightarrow \mathsf{T}(\mathbf{w}) = \mathbf{0} \Rightarrow \mathbf{w} \in \mathsf{N}(\mathsf{T})$. Thus, $\mathbf{w} \in \mathsf{N}(\mathsf{T}) \cap \mathsf{W} \Rightarrow \mathsf{N}(\mathsf{T}_\mathsf{W}) \subset \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$.

Let $\mathbf{u} \in \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$. Then, $\mathbf{u} \in \mathsf{N}(\mathsf{T}) \Rightarrow \mathsf{T}(\mathbf{u}) = \mathbf{0}$. Since $\mathbf{u} \in \mathsf{W}$ and W is T-invariant, $\mathsf{T}_\mathsf{W}(\mathbf{u}) = \mathsf{T}(\mathbf{u}) = \mathbf{0} \in \mathsf{W} \Rightarrow \mathbf{u} \in \mathsf{N}(\mathsf{T}_\mathsf{W})$. Hence, $\mathsf{N}(\mathsf{T}) \cap \mathsf{W} \subset \mathsf{N}(\mathsf{T}_\mathsf{W})$.

Therefore, $\mathsf{N}(\mathsf{T}_\mathsf{W}) = \mathsf{N}(\mathsf{T}) \cap \mathsf{W}$. Hence, proved.