MATH2211 SPRING 2022
PROBLEM SET 7 SOLUTIONS

DUE WEDNESDAY, MARCH 30, 2022 AT 11:59 PM

Problem 1. Suppose $A \in M_{m\times n}(F)$. Prove or find counterexamples:

(a) For any $B \in M_{n\times p}(F)$, $\operatorname{rank}(AB) \leq \operatorname{rank}(A)$.

(b) For any $B \in M_{p\times m}(F)$, $\operatorname{rank}(BA) \leq \operatorname{rank}(A)$.

Solution

Here is an incredibly short and concise reason why both statements are true. The image of a vector space under a matrix always has dimension less than or equal to the vector space you started with, no matter what the matrix is. Moreover, the rank of a matrix is at most the dimension of the source space.
Some notation: For any linear transformation $T\colon V \to W$ and for any subspace $U \subseteq V$, we write TU to mean the space $\{Tv : v \in U\}$. This is also the image of $T|_U$, the restriction of T to U.
Let $V = F^p$ in part (a). Then $BV = \operatorname{im} B \subseteq F^n$, thus $ABV \subseteq AF^n = \operatorname{im} A$. Therefore, $\operatorname{rank}(AB) = \dim(ABV) \leq \dim(AF^n) = \operatorname{rank} A$.
Let $V = F^m$ in part (b). Then AV has dimension $(\operatorname{rank} A)$, and so $\dim(BAV) = \operatorname{im}(B|_{AV}) \leq \dim(AV) = \operatorname{rank} A$.

Problem 2.

(a) Prove, using the exterior algebra definition of the determinant, that the determinant of an upper-triangular square matrix is the product of the diagonal entries.

Solution

Let $(a_{ij})_{1\leq i,j\leq n}$ be an $n \times n$ matrix. The wedge product of the columns of this matrix is

$$(a_{11}e_1) \wedge (a_{12}e_1 + a_{22}e_2) \wedge \cdots \wedge (a_{1n}e_1 + \cdots + a_{nn}e_n).$$

Because the first term is a scalar multiple of e_1, the e_1 coefficients in the rest of the terms do not contribute to the final expanded wedge product.

Therefore, the above n-vector is equal to

$$(a_{11}e_1) \wedge (a_{22}e_2) \wedge (a_{23}e_2 + a_{33}e_3) \wedge \cdots \wedge (a_{2n}e_2 + \cdots + a_{nn}e_n).$$

We can continue this argument until the n-vector simplifies to

$$a_{11}a_{22}\cdots a_{nn} \cdot e_1 \wedge \cdots \wedge e_n.$$

Thus the determinant of $(a_{ij})_{1\leq i,j\leq n}$ is equal to $a_{11}a_{22}\cdots a_{nn}$.
This is sort of an "unrigorous" way to do induction. But I think this is acceptable as long as it's easy to check that it can be turned into a rigorous proof by induction.

(b) You may assume that a similar proof as in part (a) shows that the same result holds for lower triangular square matrices as well.

Prove that that the elementary matrix E representing the row operation "add a multiple of row j to row i" ($i \neq j$) has determinant 1.

Solution

The given elementary matrix E has 1 along the diagonal and 0 everywhere else except for a single nonzero entry somewhere above or below the diagonal. This makes E upper or lower triangular, respectively, so part (a) shows that the determinant of E is 1.

Problem 3.

(a) Prove that if $T: V \to W$ is surjective, then ${}^tT: W^* \to V^*$ is injective.

Solution

Suppose that ${}^tT\mu = 0$ for some $\mu \in W^*$. Thus $\mu T = 0$. Saying that μT is the zero linar functional on V is the same thing as saying that $\mu T v = 0$ for all $v \in V$. Since T is surjective, we have $\{Tv : v \in V\} = W$. Thus $\mu w = 0$ for all $w \in W$. Thus μ is the zero linear functional. Therefore, ${}^tT$ is injective.

(b) Prove that if $T: V \to W$ is injective, then ${}^tT: W^* \to V^*$ is surjective.

Solution

Here's a somewhat clean proof avoiding bases. The injectivity of T implies that T is an isomorphism of V onto its image, which we call TV.
To show that ${}^tT$ is surjective, let's take an arbitrary $\lambda \in V^*$ and try to show that there exists $\mu \in W^*$ such that $\mu T = \lambda$. Let us split W into a direct sum

$W = TV \oplus Z$, hence every vector $w \in W$ is uniquely a sum $Tv + z$ for some $v \in V$ and $z \in Z$. This mapping is well defined because T is an isomorphism. Let us define $\mu\colon W \to \mathbb{R}$ by $\mu(Tv + z) = \lambda(v)$ for all $v \in V$, $z \in Z$. One easily checks that μ is a linear functional and that $\mu T v = \lambda v$ for all $v \in V$, showing that $\mu T = \lambda$.
Note: this isn't the only μ that can be chosen. We could have chosen $\mu(Tv + z) = \lambda(v) + \eta(z)$ for any linear functional η on Z. This corresponds to the fact that ${}^tT$ is not necessarily injective.

(c) Prove that every linear transformation $T\colon V \to W$ can be factored as a composition

$$V \overset{T'}{\twoheadrightarrow} U \overset{i}{\hookrightarrow} W,$$

where T' is surjective and i is injective. (Hint: Set $U = \operatorname{im} T$.)

Solution

The hint basically solves everything. Set $U = \operatorname{im} T$, set T' to be the codomain restriction of T to $\operatorname{im} T$, and let i be the inclusion $\operatorname{im} T \hookrightarrow W$. By definition of $\operatorname{im} T$, every vector in $\operatorname{im} T$ is equal to Tv (which is also $T'v$) for some $v \in V$, so T' is surjective. Moreover, i is clearly injective because it is an inclusion map.

(d) Recall that the rank of a linear transformation $T\colon V \to W$ is defined to be $\dim \operatorname{im} T$. Use parts (a), (b), and (c) to prove that $\operatorname{rank}(T) = \operatorname{rank}({}^tT)$.

Solution

The only ingredient needed to finish the proof is that the dimension of the middle term in any surjective-injective factorization (i.e. some other factorization besides the one with $U = \operatorname{im} T$ as in part (c)) depends only on T and not on the factorization. This is because $\dim U = \dim(iU) = \dim(\operatorname{im} T)$ by the fact that $T = iT'$.
Now let us prove that $\operatorname{rank}(T) = \operatorname{rank}({}^tT)$. We can observe two factorizations of ${}^tT$ into surjective and injective maps. The first one is

$$W^* \twoheadrightarrow \operatorname{im}({}^tT) \hookrightarrow V^*$$

obtained by applying part (c) to ${}^tT$. The second one is

$$W^* \overset{{}^ti}{\twoheadrightarrow} (\operatorname{im} T)^* \overset{{}^tT'}{\hookrightarrow} V^*$$

obtained by dualizing

$$V \overset{T'}{\twoheadrightarrow} \operatorname{im} T \overset{i}{\hookrightarrow} W$$

and using parts (a) and (b) to see that ${}^t i$ is surjective and ${}^t T'$ is injective. Therefore, using the ingredient in the first paragraph, we find that $\dim((\operatorname{im} T)^*) = \dim \operatorname{im}({}^t T)$. But we already know that $\dim((\operatorname{im} T)^*) = \dim(\operatorname{im} T)$, so therefore have proved that $\operatorname{rank}(T) = \operatorname{rank}({}^t T)$.

Problem 4. Let V be a vector space. Given a nonzero element $v \in V$ and a nonzero linear functional $\lambda \colon V \to \mathbb{R}$, we can make a linear transformation $T_{v,\lambda} \colon V \to V$ by sending $x \in V$ to $\lambda(x) \cdot v$. Prove that $T_{v,\lambda}$ has rank 1, and prove that every rank 1 linear transformation from V to V is equal to $T_{v,\lambda}$ for some $0 \neq v \in V$ and $0 \neq \lambda \colon V \to \mathbb{R}$.

Hint for the second part: Problem 3(c) is very helpful.

Solution

The linear transformation $T_{v,\lambda}$ has rank 1 because the image of $T_{v,\lambda}$ consists of scalar multiples of v (and these scalar multiples are nonzero because λ is not the zero linear functional.)
Let T now be a rank 1 matrix. Therefore, the image of T is a 1-dimensional subspace of V. Let $\{v\}$ be any basis of this subspace. For each $w \in V$, we know that Tw is some multiple of v. Let us define a function $\lambda(w) \in F$ defined by $Tw = \lambda(w)v$ for all $w \in V$. Now one easily checks that in fact λ depends linearly on w, i.e. is a linear transformation, so λ is a linear functional. Therefore, $T = T_{v,\lambda}$ for this v and this λ.

Problem 5. Let $\frac{d}{dx} \colon P_2(\mathbb{R}) \to P_2(\mathbb{R})$ be the derivative operator on the space of polynomials of degree at most 2 over $\mathbb{R}$.

(a) Prove that ${}^t(\frac{d}{dx}) \colon P_2(\mathbb{R})^* \to P_2(\mathbb{R})^*$ has a 1-dimensional kernel.

Solution

Using Problem 3 we find that $2 = \operatorname{rank} \frac{d}{dx} = \operatorname{rank}({}^t(\frac{d}{dx}))$. Thus the nullity of ${}^t(\frac{d}{dx})$ is $3 - 2 = 1$.

(b) Prove that $\ker({}^t(\frac{d}{dx}))$ is spanned by the functional taking a polynomial to its x^2 coefficient. Can you provide a plain English interpretation of what this is saying?

Solution

This can be done with matrices, but here is a neater proof. First, write $[x^2]$ as the linear functional $P_2(\mathbb{R}) \to F; p \mapsto$ the coefficient of x^2 in p. This just lets us have a convenient name to call this functional.

For any polynomial $p \in P_2(\mathbb{R})$, we have $({}^t(\frac{d}{dx})[x^2])p = [x^2]\frac{dp}{dx}$. Now, the derivative of a polynomial of degree 2 has degree 1, so $[x^2]\frac{dp}{dx} = 0$. Hence $[x^2] \in \ker({}^t(\frac{d}{dx}))$ for all $p \in P_2(\mathbb{R})$. By part (a), the kernel is one-dimensional, so we are done.

The English interpretation of the fact that $[x^2]$ spans the kernel of the transpose of the derivative operator is that the derivative of any polynomial in $P_2(\mathbb{R})$ always has x^2 coefficient equal to 0. This is basically what we used in the proof above, but is a new interpretation if you solved this problem with matrices.

It's quite interesting to see that analyzing $\ker \frac{d}{dt}$ and $\ker({}^t(\frac{d}{dt}))$ show two very different facts about the derivative!