2.2: 8, 10, 11

2.3: 11, 12, 13

2.2:

8. Let V be an n-dimensional vector space with an ordered basis β. Define $\mathsf{T} : \mathsf{V} \to \mathbb{F}^n$ by $\mathsf{T}(x) = [x]_\beta$. Prove that T is linear.

Proof. The transformation T is linear iff $\mathsf{T}(ax+y) = a\mathsf{T}(x) + \mathsf{T}(y)$ for vectors $x, y \in \mathsf{V}$ and scalar $a \in \mathbb{F}$.

So $\mathsf{T}(ax+y) = [ax+y]_\beta$, which is the coordinate vector representation of $ax+y$ in terms of the ordered basis given by β.

We can represent the vector given by $[ax+y]_\beta$ as a linear combination of other vectors relative to the same basis, namely $[x]_\beta$ and $[y]_\beta$. It is true that $a\,[x]_\beta + [y]_\beta = [ax]_\beta + [y]_\beta = [ax+y]_\beta$, so then we can express $a\,[x]_\beta + [y]_\beta$ as $a\mathsf{T}(x) + \mathsf{T}(y)$. Therefore T is linear. $\square$

10. For every vector v_j in the ordered basis β, we have that $\mathsf{T}(v_j) = v_j + v_{j-1}$. The matrix $[\mathsf{T}]_\beta$ can be expressed as $\left([\mathsf{T}(v_1)]_\beta\,[\mathsf{T}(v_2)]_\beta \cdots [\mathsf{T}(v_n)]_\beta\right)$, where each $[\mathsf{T}(v_i)]_\beta$ is a column.

The first column is unique since $v_{1-1} = v_0 = \vec{0}$, so we have that

$$[\mathsf{T}(v_1)]_\beta = v_1 + v_0 = v_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

However, the other columns are in the form where there is an extra '1' in the position above the '1' meant to carry v_j to itself. So for example,

$$[\mathsf{T}(v_2)]_\beta = v_2 + v_1 = \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, [\mathsf{T}(v_3)]_\beta = v_3 + v_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix},$$

$$[\mathsf{T}(v_{n-1})]_\beta = v_{n-1} + v_{n-2} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 1 \\ 0 \end{pmatrix}, \text{ and } [\mathsf{T}(v_n)]_\beta = v_n + v_{n-1} = \begin{pmatrix} 0 \\ \vdots \\ 1 \\ 1 \end{pmatrix}$$

Thus by construction

$$[\mathsf{T}]_\beta = \Big([\mathsf{T}(v_1)]_\beta \, [\mathsf{T}(v_2)]_\beta \cdots [\mathsf{T}(v_n)]_\beta\Big) = \begin{pmatrix} 1 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 1 & \cdots & 0 & 0 \\ 0 & 0 & 1 & \ddots & \vdots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & 1 \\ 0 & 0 & \cdots & 0 & 0 & 1 \end{pmatrix}.$$

11. Let V be an n-dimensional vector space, and let $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be a linear transformation. Suppose that W is a T-invariant subspace of V having dimension k. Show that there is a basis β for V such that $[\mathsf{T}]_\beta$ has the form

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

where A is a $k \times k$ matrix and O is the $(n-k) \times k$ zero matrix.

Proof. Consider a basis for W given by $\{w_1, w_2, \ldots, w_k\}$. Then by the replacement theorem we can extend this basis for W into a basis for V by adding $n-k$ many linearly independent vectors into the basis, resulting in the ordered basis β given by $\{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$.

Since W is T-invariant, the construction of $[\mathsf{T}]_\beta$ is greatly simplified, as the transformations $\mathsf{T}(w_i)$ will always be in the form

$$\mathsf{T}(w_i) = \begin{pmatrix} c_{1i} \\ c_{2i} \\ \vdots \\ c_{ki} \\ 0_1 \\ 0_2 \\ \vdots \\ 0_{(n-k)} \end{pmatrix},$$

where because $\mathsf{T}(w_i) \in \mathsf{W}$, and each v_i are not in W, we have that the corresponding entries for each v_i in the coordinate vector must be zero (denoted by 0_i). Consider the construction for the first part of $[\mathsf{T}]_\beta$, given by

$$\Big([\mathsf{T}(w_1)]_\beta\,[\mathsf{T}(w_2)]_\beta \ldots [\mathsf{T}(w_k)]_\beta\Big) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1k} \\ a_{21} & a_{22} & \cdots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \cdots & a_{kk} \\ 0_1 & 0_1 & \cdots & 0_1 \\ 0_2 & 0_2 & \cdots & 0_2 \\ \vdots & \vdots & \cdots & \vdots \\ 0_{(n-k)} & 0_{(n-k)} & \cdots & 0_{(n-k)} \end{pmatrix} = \begin{pmatrix} A \\ O \end{pmatrix}.$$

Then for the remaining vectors v_i in β, we construct $\Big([\mathsf{T}(v_1)]_\beta\,[\mathsf{T}(v_2)]_\beta \ldots [\mathsf{T}(v_{n-k})]_\beta\Big)$ and adjoin it to the previous matrix. Here the matrix will not have any special conditions since naturally each v_i can be expressed as a linear combination of vectors in β.

$$\Big([\mathsf{T}(v_1)]_\beta\,[\mathsf{T}(v_2)]_\beta \ldots [\mathsf{T}(v_{n-k})]_\beta\Big) = \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1(n-k)} \\ b_{21} & b_{22} & \cdots & b_{2(n-k)} \\ \vdots & \vdots & \cdots & \vdots \\ b_{k1} & b_{k2} & \cdots & b_{k(n-k)} \\ c_{11} & c_{12} & \cdots & c_{1(n-k)} \\ c_{21} & c_{22} & \cdots & c_{2(n-k)} \\ \vdots & \vdots & \ddots & \vdots \\ c_{(n-k)1} & c_{(n-k)2} & \cdots & c_{(n-k)(n-k)} \end{pmatrix} = \begin{pmatrix} B \\ C \end{pmatrix}.$$

Thus

$$[\mathsf{T}]_\beta = \Big([\mathsf{T}(w_1)]_\beta\,[\mathsf{T}(w_2)]_\beta \ldots [\mathsf{T}(w_k)]_\beta\,[\mathsf{T}(v_1)]_\beta\,[\mathsf{T}(v_2)]_\beta \ldots [\mathsf{T}(v_{n-k})]_\beta\Big) = \begin{pmatrix} A & B \\ O & C \end{pmatrix},$$

so there is a basis $\beta = \{w_1, w_2, \ldots, w_k, v_1, v_2, \ldots, v_{n-k}\}$ for V that gives $[\mathsf{T}]_\beta$ this form. □

2.3:

11. Let V be a vector space, and let $\mathsf{T} : \mathsf{V} \to \mathsf{V}$ be linear. Prove that $\mathsf{T}^2 = \mathsf{T}_0$ if and only if $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$.

Proof. Forwards direction. Suppose $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$. Then for all $v \in \mathsf{V}$, $\mathsf{T}(v) \in \mathsf{R}(\mathsf{T})$ whenever $\mathsf{T}(v) \neq \vec{0}$ (in which case $\mathsf{T}(\vec{0}) = \vec{0}$ and nothing needs to be done). Then by assumption $\mathsf{T}(v) \in \mathsf{N}(\mathsf{T})$, so that $\mathsf{T}(\mathsf{T}(v)) = \vec{0} = \mathsf{T}^2(v)$. Since this holds for all $v \in \mathsf{V}$, T^2 must be the zero transformation.

Converse. Suppose $\mathsf{T}^2 = \mathsf{T}_0$. Then for all $v \in \mathsf{V}$, $\mathsf{T}^2(v) = \vec{0} = \mathsf{T}(\mathsf{T}(v))$. We know that for $v \neq \vec{0}$ (the trivial case), $\mathsf{T}(v) \in \mathsf{R}(\mathsf{T})$, so then if all such elements of $\mathsf{R}(T)$ map into the zero vector (by $\mathsf{T}(\mathsf{T}(v)) = \vec{0}$), then all elements of $\mathsf{R}(\mathsf{T})$ must belong to the null space. Therefore $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$.

Hence $\mathsf{T}^2 = \mathsf{T}_0$ if and only if $\mathsf{R}(\mathsf{T}) \subseteq \mathsf{N}(\mathsf{T})$. □

12. Let V, W, and Z be vector spaces, and let $\mathsf{T}:\mathsf{V}\to\mathsf{W}$ and $\mathsf{U}:\mathsf{W}\to\mathsf{Z}$ be linear.

(a) Prove that if UT is one-to-one, then T is one-to-one. Must U also be one-to-one?

Proof. Suppose UT is one-to-one. Then suppose by way of contradiction that T is not one-to-one, that is, there are two distinct vectors $x, y \in \mathsf{V}$ such that $\mathsf{T}(x)=\mathsf{T}(y)$.

Consider $\mathsf{UT}(x)=\mathsf{U}(\mathsf{T}(x))=\mathsf{U}(\mathsf{T}(y))=\mathsf{UT}(y)$,for two distinct vectors x, y. So UT is not injective as assumed, which is a contradiction. Therefore T must be injective. □

No, U need not be one-to-one, since if $\mathsf{T}(x)=\mathsf{T}(y)$, then $\mathsf{U}(\mathsf{T}(x))=\mathsf{U}(\mathsf{T}(y))$. So then $\mathsf{UT}(x)=\mathsf{UT}(y)$, which means $x=y$, regardless of what U is (provided UT is injective).

(b) Prove that if UT is onto, then U is onto. Must T also be onto?

Proof. Suppose UT is onto. Then suppose by way of contradiction that U is not onto, that is, there exists a vector z in Z that is not the image of any vector in W.

Then it is impossible for UT to be onto because for any vector $v \in \mathsf{V}$, $\mathsf{T}(v)\in\mathsf{W}$ but no such vector $\mathsf{T}(v)$ can be the preimage of z under U. This is in contradiction to the assumption that UT is onto and so we must have that U is onto. □

No, T need not be onto, because for any T we still know that UT will be onto. Then this means that there exists a $v \in \mathsf{V}$ such that for every $z \in \mathsf{Z}$, $z=\mathsf{UT}(v)=\mathsf{U}(\mathsf{T}(v))$, where $\mathsf{T}(v)$ is guaranteed to map to z due to U being surjective, so T need not be onto.

(c) Prove that if U and T are bijective, then UT is also.

Proof. Suppose U and T are bijective. Then show that UT is injective and surjective.

Injectivity. Suppose x, y are distinct vectors in V. Then $\mathsf{UT}(x)=\mathsf{U}(\mathsf{T}(x))$ and $\mathsf{UT}(y)=\mathsf{U}(\mathsf{T}(y))$, where $\mathsf{T}(x)\neq\mathsf{T}(y)$ due to the injectivity of T. Then similarly by injectivity of U, $\mathsf{U}(\mathsf{T}(x))\neq\mathsf{U}(\mathsf{T}(y))$ and so $\mathsf{UT}(x)\neq\mathsf{UT}(y)$, which implies UT is injective.

Surjectivity. We wish to show that for all $z \in \mathsf{Z}$, there is an element $v \in \mathsf{V}$ such that $\mathsf{UT}(v)=z$. Since U is surjective, there exists a vector $w \in \mathsf{W}$ such that $\mathsf{U}(w)=z$. Then similarly since T is surjective, then there exists a vector $v \in \mathsf{V}$ such that $\mathsf{T}(v)=w$. Then $\mathsf{U}(\mathsf{T}(v))=\mathsf{UT}(v)=z$ for all $z \in \mathsf{Z}$. Hence UT is surjective.

Therefore UT is bijective. □

13. Let A and B be $n \times n$ matrices. Prove that $\operatorname{tr}(AB)=\operatorname{tr}(BA)$ and $\operatorname{tr}(A)=\operatorname{tr}\left(A^t\right)$.

Proof. (1) Each entry of AB is given by $(AB)_{ij}=\sum_{k=1}^{n} A_{ik}B_{kj}$, by rules of matrix multiplication. Then the elements on the diagonal are $(AB)_{ii} = \sum_{k=1}^{n} A_{ik}B_{ki}$, and so

$$\operatorname{tr}(AB) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} A_{ik}B_{ki} \right).$$

But by swapping the order of summation and commuting the term in the summand we can show that

$$\operatorname{tr}(AB) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} B_{ki}A_{ik} \right).$$

However, entries of BA take on the form $(BA)_{ij} = \sum_{k=1}^{n} B_{ik}A_{kj}$, and the entries in the diagonal are $(BA)_{ii} = \sum_{k=1}^{n} B_{ik}A_{ki}$. So the trace of BA can be expressed as

$$\operatorname{tr}(BA) = \sum_{i=1}^{n} \left(\sum_{k=1}^{n} B_{ik}A_{ki} \right),$$

but because summation variables are dummy variables we may replace k by i and vice versa to find that

$$\operatorname{tr}(BA) = \sum_{k=1}^{n} \left(\sum_{i=1}^{n} B_{ki}A_{ik} \right) = \operatorname{tr}(AB).$$

Hence $\operatorname{tr}(AB) = \operatorname{tr}(BA)$. □

Proof. (2) The trace of A is $\sum_{i=1}^{n} A_{ii}$, since elements of A in position ij are A_{ij} and elements on the diagonal are where $i = j$. Similarly elements of A^t are $\left(A^t\right)_{ij} = A_{ji}$, and the elements on the diagonal are also where $i = j$, so the trace of A^t is $\sum_{i=1}^{n} \left(A^t\right)_{ii} = \sum_{i=1}^{n} A_{ii}$

Hence $\operatorname{tr}(A) = \operatorname{tr}\left(A^t\right)$. □