线性变换

Linear Transformations

Baobin Li

Email:libb@ucas.ac.cn

School of Computer and Control Engineering, UCAS

Introduction

- The connection between linear functions and matrices is at the heart of our subject.
- It is now time to formalize the connections between matrices, vector spaces, and linear functions defined on vector spaces.

Linear Transformations

Let $\mathcal{U}$ and $\mathcal{V}$ be vector spaces over a field $\mathcal{F}$ ($\Re$ or $\mathcal{C}$ for us).

- A ***linear transformation*** from $\mathcal{U}$ into $\mathcal{V}$ is defined to be a linear function $\mathbf{T}$ mapping $\mathcal{U}$ into $\mathcal{V}$. That is,

  $$\mathbf{T}(\mathbf{x}+\mathbf{y}) = \mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y}) \quad \text{and} \quad \mathbf{T}(\alpha\mathbf{x}) = \alpha\mathbf{T}(\mathbf{x})$$

  or, equivalently,

  $$\mathbf{T}(\alpha\mathbf{x}+\mathbf{y}) = \alpha\mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y}) \quad \text{for all } \mathbf{x},\mathbf{y} \in \mathcal{U},\ \alpha \in \mathcal{F}.$$

- A ***linear operator*** on $\mathcal{U}$ is defined to be a linear transformation from $\mathcal{U}$ into itself—i.e., a linear function mapping $\mathcal{U}$ back into $\mathcal{U}$.

Some Example

- The function $\mathbf{0}(x) = \mathbf{0}$ that maps all vectors in a space $\mathcal{U}$ to the zero vector in another space $\mathcal{V}$ is a linear transformation from $\mathcal{U}$ into $\mathcal{V}$, and, not surprisingly, it is called the **zero transformation**.
- The function $\mathbf{I}(x) = x$ that maps every vector from a space $\mathcal{U}$ back to itself is a linear operator on $\mathcal{U}$. $\mathbf{I}$ is called the **identity operator** on $\mathcal{U}$.
- For $\mathbf{A} \in \Re^{m\times n}$ and $x \in \Re^{n\times 1}$, the function $\mathbf{T}(x) = \mathbf{A}x$ is a linear transformation from $\Re^n$ into $\Re^m$ because matrix multiplication satisfies $\mathbf{A}(\alpha x + y) = \alpha \mathbf{A}x + \mathbf{A}y$. $\mathbf{T}$ is a linear operator on $\Re^n$ if A is $n \times n$.
- If $\mathcal{W}$ is the vector space of all functions from $\Re$ to $\Re$, and if $\mathcal{V}$ is the space of all differentiable functions from $\Re$ to $\Re$, then the mapping $\mathbf{D}(f) = df/dx$ is a linear transformation from $\mathcal{V}$ into $\mathcal{W}$ because

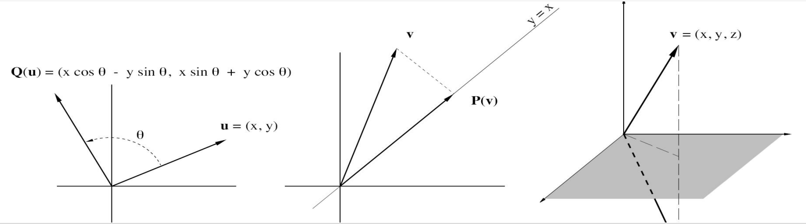
$$\frac{d(\alpha f + g)}{dx} = \alpha \frac{df}{dx} + \frac{dg}{dx}.$$

- If $\mathcal{V}$ is the space of all continuous functions from $\mathfrak{R}$ into $\mathfrak{R}$, then the mapping defined by $\mathbf{T}(f) = \int_0^x f(t)dt$ is a linear operator on $\mathcal{V}$ because

$$\int_0^x [\alpha f(t) + g(t)]dt = \alpha \int_0^x f(t)dt + \int_0^x g(t)dt.$$

- The rotator $\mathbf{Q}$ that rotates vectors $\mathbf{u}$ in $\mathfrak{R}^2$ counterclockwise through an angle θ, is a linear operator on $\mathfrak{R}^2$ because the "action" of $\mathbf{Q}$ on $\mathbf{u}$ can be described by matrix multiplication in the sense that the coordinates of the rotated vector $\mathbf{Q}(\mathbf{u})$ are given by

$$\mathbf{Q}(\mathbf{u}) = \begin{pmatrix} xcos\theta - ysin\theta \\ xsin\theta + ycos\theta \end{pmatrix} = \begin{pmatrix} cos\theta & -sin\theta \\ sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}.$$



- The projector $\mathbf{P}$ that maps $\mathbf{v} = (x, y, z) \in \Re^3$ to its orthogonal projection $(x, y, 0)$ in the xy -plane, is a linear operator on $\Re^3$ because if $\mathbf{u} = (u1, u2, u3)$ and $\mathbf{v} = (v1, v2, v3)$, then

$$\mathbf{P}(\alpha\mathbf{u} + \mathbf{v}) = (\alpha u1 + v1, \alpha u2 + v2, 0) = \alpha\mathbf{P}(\mathbf{u}) + \mathbf{P}(\mathbf{v}).$$

- The reflector $\mathbf{R}$ that maps each vector $v = (x, y, z) \in \Re^3$ to its reflection $\mathbf{R}(v) = (x, y, -z)$ about the xy -planeis a linear operator on $\Re^3$.
- Can all linear transformations be represented by matrices?
- Linear transformations on finite-dimensional spaces will always have matrix representations.
- To see why, the concept of coordinates in higher dimensions must first be understood.
- Recall that if $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n\}$ is a basis for a vector space $\mathcal{U}$, then $\mathbf{v} = \alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \cdots + \alpha_n\mathbf{u}_n$.
- α_i's are uniquely determined by $\mathbf{v}$, which are called the **coordinates of $\mathbf{v}$ with respect to $\mathcal{B}$** . Denote $[\mathbf{v}]_{\mathcal{B}} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T$.

- From now on, $\mathcal{S} = \{\mathbf{e}_1, \mathbf{e}_2, \cdots, \mathbf{e}_n\}$ will denote the standard basis of unit vectors for $\mathfrak{R}^n$.
- Linear transformations possess coordinates in the same way vectors do because linear transformations from $\mathcal{U}$ to $\mathcal{V}$ also form a vector space.

Space of Linear Transformations

- For each pair of vector spaces $\mathcal{U}$ and $\mathcal{V}$ over $\mathcal{F}$, the set $\mathcal{L}(\mathcal{U}, \mathcal{V})$ of all linear transformations from $\mathcal{U}$ to $\mathcal{V}$ is a vector space over $\mathcal{F}$.
- Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively, and let $\mathbf{B}_{ji}$ be the linear transformation from $\mathcal{U}$ into $\mathcal{V}$ defined by $\mathbf{B}_{ji}(\mathbf{u}) = \xi_j \mathbf{v}_i$, where $(\xi_1, \xi_2, \ldots, \xi_n)^T = [\mathbf{u}]_{\mathcal{B}}$. That is, pick off the j^{th} coordinate of $\mathbf{u}$, and attach it to $\mathbf{v}_i$.
  - ▷ $\mathcal{B}_{\mathcal{L}} = \{\mathbf{B}_{ji}\}_{j=1\ldots n}^{i=1\ldots m}$ is a basis for $\mathcal{L}(\mathcal{U}, \mathcal{V})$.
  - ▷ $\dim \mathcal{L}(\mathcal{U}, \mathcal{V}) = (\dim \mathcal{U})(\dim \mathcal{V})$.

- Prove $\mathcal{B}_{\mathcal{L}}$ is a basis by demonstrating that it is a linearly independent spanning set for $\mathcal{L}(\mathcal{U},\mathcal{V})$. To establish linear independence,

suppose $\sum_{j,i}\eta_{ji}\mathbf{B}_{ji}=\mathbf{0}$ for scalars η_{ji}, and observe that for each $\mathbf{u}_k\in\mathcal{B}$,

$$\mathbf{B}_{ji}(\mathbf{u}_k)=\begin{cases}\mathbf{v}_i & \text{if } j=k\\ \mathbf{0} & \text{if } j\neq k\end{cases} \Longrightarrow \mathbf{0}=\Big(\sum_{j,i}\eta_{ji}\mathbf{B}_{ji}\Big)(\mathbf{u}_k)=\sum_{j,i}\eta_{ji}\mathbf{B}_{ji}(\mathbf{u}_k)=\sum_{i=1}^{m}\eta_{ki}\mathbf{v}_i.$$

For each k, the independence of $\mathcal{B}'$ implies that $\eta_{ki}=0$ for each i, and thus $\mathcal{B}_{\mathcal{L}}$ is linearly independent. To see that $\mathcal{B}_{\mathcal{L}}$ spans $\mathcal{L}(\mathcal{U},\mathcal{V})$, let $\mathbf{T}\in\mathcal{L}(\mathcal{U},\mathcal{V})$, and determine the action of $\mathbf{T}$ on any $\mathbf{u}\in\mathcal{U}$ by using $\mathbf{u}=\sum_{j=1}^{n}\xi_j\mathbf{u}_j$ and $\mathbf{T}(\mathbf{u}_j)=\sum_{i=1}^{m}\alpha_{ij}\mathbf{v}_i$ to write

$$\mathbf{T}(\mathbf{u})=\mathbf{T}\left(\sum_{j=1}^{n}\xi_j\mathbf{u}_j\right)=\sum_{j=1}^{n}\xi_j\mathbf{T}(\mathbf{u}_j)=\sum_{j=1}^{n}\xi_j\sum_{i=1}^{m}\alpha_{ij}\mathbf{v}_i$$
$$=\sum_{i,j}\alpha_{ij}\xi_j\mathbf{v}_i=\sum_{i,j}\alpha_{ij}\mathbf{B}_{ji}(\mathbf{u}).$$

This holds for all $\mathbf{u}\in\mathcal{U}$, so $\mathbf{T}=\sum_{i,j}\alpha_{ij}\mathbf{B}_{ji}$, and thus $\mathcal{B}_{\mathcal{L}}$ spans $\mathcal{L}(\mathcal{U},\mathcal{V})$.

Coordinate Matrix Representations

Let $\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ and $\mathcal{B}' = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_m\}$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. The ***coordinate matrix*** of $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ with respect to the pair $(\mathcal{B}, \mathcal{B}')$ is defined to be the $m \times n$ matrix

$$[\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \left([\mathbf{T}(\mathbf{u}_1)]_{\mathcal{B}'} \,\middle|\, [\mathbf{T}(\mathbf{u}_2)]_{\mathcal{B}'} \,\middle|\, \cdots \,\middle|\, [\mathbf{T}(\mathbf{u}_n)]_{\mathcal{B}'} \right).$$

In other words, if $\mathbf{T}(\mathbf{u}_j) = \alpha_{1j}\mathbf{v}_1 + \alpha_{2j}\mathbf{v}_2 + \cdots + \alpha_{mj}\mathbf{v}_m$, then

$$[\mathbf{T}(\mathbf{u}_j)]_{\mathcal{B}'} = \begin{pmatrix} \alpha_{1j} \\ \alpha_{2j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \text{ and } [\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}.$$

When $\mathbf{T}$ is a linear operator on $\mathcal{U}$, and when there is only one basis involved, $[\mathbf{T}]_{\mathcal{B}}$ is used in place of $[\mathbf{T}]_{\mathcal{B}\mathcal{B}}$ to denote the (necessarily square) coordinate matrix of $\mathbf{T}$ with respect to $\mathcal{B}$.

- **Example:** If $\mathbf{P}$ is the projector, determine the coordinate matrix $[\mathbf{P}]_{\mathfrak{B}}$ with respect to the basis

$$\mathcal{B} = \left\{ \mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \mathbf{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \mathbf{u}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \right\}.$$

- **Solution:** According to the above results, the j^{th} column in $[\mathbf{P}]_{\mathfrak{B}}$ is $[\mathbf{P}(\mathbf{u}_j)]_{\mathfrak{B}}$. Therefore,

$$\mathbf{P}(\mathbf{u}_1) = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1\mathbf{u}_1 + 1\mathbf{u}_2 - 1\mathbf{u}_3 \implies [\mathbf{P}(\mathbf{u}_1)]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_2) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 0\mathbf{u}_1 + 3\mathbf{u}_2 - 2\mathbf{u}_3 \implies [\mathbf{P}(\mathbf{u}_2)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix},$$

$$\mathbf{P}(\mathbf{u}_3) = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} = 0\mathbf{u}_1 + 3\mathbf{u}_2 - 2\mathbf{u}_3 \implies [\mathbf{P}(\mathbf{u}_3)]_{\mathcal{B}} = \begin{pmatrix} 0 \\ 3 \\ -2 \end{pmatrix},$$

$$\text{so that } [\mathbf{P}]_{\mathcal{B}} = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 3 & 3 \\ -1 & -2 & -2 \end{pmatrix}.$$

- At the heart of linear algebra is the realization that the theory of finite dimensional linear transformations is essentially the same as the theory of matrices.
- This is due primarily to the fundamental fact that the action of a linear transformation $\mathbf{T}$ on a vector $\mathbf{u}$ is precisely matrix multiplication between the coordinates of $\mathbf{T}$ and the coordinates of $\mathbf{u}$.

Action as Matrix Multiplication

Let $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and let $\mathcal{B}$ and $\mathcal{B}'$ be bases for $\mathcal{U}$ and $\mathcal{V}$, respectively. For each $\mathbf{u} \in \mathcal{U}$, the action of $\mathbf{T}$ on $\mathbf{u}$ is given by matrix multiplication between their coordinates in the sense that

$$[\mathbf{T}(\mathbf{u})]_{\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}\mathcal{B}'}[\mathbf{u}]_{\mathcal{B}}.$$

- **Example:** Show how the action of the operator $\mathbf{D}\,(p(t)) = dp/dt$ on the space $\mathcal{P}_3$ of polynomials of degree three or less is given by matrix multiplication.

Solution: The coordinate matrix of $\mathbf{D}$ with respect to the basis $\mathcal{B} = \{1, t, t^2, t^3\}$ is

$$[\mathbf{D}]_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

If $\mathbf{p} = p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \alpha_3 t^3$, then $\mathbf{D}(\mathbf{p}) = \alpha_1 + 2\alpha_2 t + 3\alpha_3 t^2$ so that

$$[\mathbf{p}]_{\mathcal{B}} = \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} \quad \text{and} \quad [\mathbf{D}(\mathbf{p})]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ 3\alpha_3 \\ 0 \end{pmatrix}.$$

The action of $\mathbf{D}$ is accomplished by means of matrix multiplication because

$$[\mathbf{D}(\mathbf{p})]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ 2\alpha_2 \\ 3\alpha_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{pmatrix} = [\mathbf{D}]_{\mathcal{B}}[\mathbf{p}]_{\mathcal{B}}.$$

- For $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $\mathbf{L} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, the composition of $\mathbf{L}$ with $\mathbf{T}$ is defined to be the function $\mathbf{C} : \mathcal{U} \to \mathcal{W}$ such that $\mathbf{C}(x) = \mathbf{L}\left(\mathbf{T}(x)\right)$.
- This composition denoted by $\mathbf{C}(x) = \mathbf{LT}$, is also a linear transformation because

$$\begin{aligned} \mathbf{C}(\alpha\mathbf{x} + \mathbf{y}) &= \mathbf{L}\left(\mathbf{T}(\alpha\mathbf{x} + \mathbf{y})\right) = \mathbf{L}\left(\alpha\mathbf{T}(\mathbf{x}) + \mathbf{T}(\mathbf{y})\right) \\ &= \alpha\mathbf{L}\left(\mathbf{T}(\mathbf{x})\right) + \mathbf{L}\left(\mathbf{T}(\mathbf{y})\right) = \alpha\mathbf{C}(\mathbf{x}) + \mathbf{C}(\mathbf{y}). \end{aligned}$$

- If $\mathfrak{B}$, $\mathfrak{B}'$ and $\mathfrak{B}''$ are bases for $\mathcal{U}$, $\mathcal{V}$ and $\mathcal{W}$, respectively, then $\mathbf{C}$ must have a coordinate matrix representation with respect to $(\mathfrak{B}, \mathfrak{B}'')$.
- So it's only natural to ask how $[\mathbf{C}]_{\mathfrak{B}\mathfrak{B}''}$ is related to $[\mathbf{L}]_{\mathfrak{B}'\mathfrak{B}''}$ and $[\mathbf{T}]_{\mathfrak{B}\mathfrak{B}'}$:

$$[\mathbf{C}]_{\mathfrak{B}\mathfrak{B}''} = [\mathbf{L}]_{\mathfrak{B}'\mathfrak{B}''}[\mathbf{T}]_{\mathfrak{B}\mathfrak{B}'}.$$

Connections with Matrix Algebra

- If $\mathbf{T}, \mathbf{L} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$, and if $\mathcal{B}$ and $\mathcal{B}'$ are bases for $\mathcal{U}$ and $\mathcal{V}$, then
  - ▷ $[\alpha\mathbf{T}]_{\mathcal{B}\mathcal{B}'} = \alpha[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$ for scalars α,
  - ▷ $[\mathbf{T} + \mathbf{L}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}\mathcal{B}'} + [\mathbf{L}]_{\mathcal{B}\mathcal{B}'}$.
- If $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{V})$ and $\mathbf{L} \in \mathcal{L}(\mathcal{V}, \mathcal{W})$, and if $\mathcal{B}$, $\mathcal{B}'$, and $\mathcal{B}''$ are bases for $\mathcal{U}$, $\mathcal{V}$, and $\mathcal{W}$, respectively, then $\mathbf{LT} \in \mathcal{L}(\mathcal{U}, \mathcal{W})$, and
  - ▷ $[\mathbf{LT}]_{\mathcal{B}\mathcal{B}''} = [\mathbf{L}]_{\mathcal{B}'\mathcal{B}''}[\mathbf{T}]_{\mathcal{B}\mathcal{B}'}$.
- If $\mathbf{T} \in \mathcal{L}(\mathcal{U}, \mathcal{U})$ is invertible in the sense that $\mathbf{TT}^{-1} = \mathbf{T}^{-1}\mathbf{T} = \mathbf{I}$ for some $\mathbf{T}^{-1} \in \mathcal{L}(\mathcal{U}, \mathcal{U})$, then for every basis $\mathcal{B}$ of $\mathcal{U}$,
  - ▷ $[\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1}$.

■ From the composition $\mathbf{C} = \mathbf{LT}$ of the two linear transformations $\mathbf{T} : \mathfrak{R}^3 \to \mathfrak{R}^2$ and $\mathbf{L} : \mathfrak{R}^2 \to \mathfrak{R}^2$ defined by

$$\mathbf{T}(x, y, z) = (x + y, y - z) \; and \; \mathbf{L}(u, v) = (2u - v, u),$$

using the standard bases for $\mathfrak{R}^2$ and $\mathfrak{R}^3$ to verify the above results.

Change of Basis and Similarity

- By their nature, coordinate matrix representations are basis dependent.
- It's desirable to study linear transformations without reference to particular bases because some bases may force a coordinate matrix representation to exhibit special properties that are not present in the coordinate matrix relative to other bases.
- It's necessary to somehow identify properties of coordinate matrices that are invariant among all bases.
- These are properties intrinsic to the transformation itself.
- The following discussion is limited to a single finite-dimensional space $\mathcal{V}$ and to linear operators on $\mathcal{V}$.
- Begin by examining how the coordinates of $\mathbf{v} \in \mathcal{V}$ change as the basis for $\mathcal{V}$ changes.
- Consider two different bases

$$\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_n\} \quad and \quad \mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_n\}.$$

- Throughout this section $\mathbf{T}$ will denote the linear operator such that

$$\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i \quad i = 1, 2, \cdots, n.$$

- $\mathbf{T}$ is called the **change of basis operator**.

Changing Vector Coordinates

Let $\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n\}$ and $\mathcal{B}' = \{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_n\}$ be bases for $\mathcal{V}$, and let $\mathbf{T}$ and $\mathbf{P}$ be the associated change of basis operator and change of basis matrix, respectively—i.e., $\mathbf{T}(\mathbf{y}_i) = \mathbf{x}_i$, for each i, and

$$\mathbf{P} = [\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = \Big([\mathbf{x}_1]_{\mathcal{B}'} \Big| [\mathbf{x}_2]_{\mathcal{B}'} \Big| \cdots \Big| [\mathbf{x}_n]_{\mathcal{B}'} \Big).$$

- $[\mathbf{v}]_{\mathcal{B}'} = \mathbf{P}[\mathbf{v}]_{\mathcal{B}}$ for all $\mathbf{v} \in \mathcal{V}$.
- $\mathbf{P}$ is nonsingular.

- To see this, observe that

$$\mathbf{x}_i = \sum_{j=1}^{n} \alpha_j \mathbf{y}_j \implies \mathbf{T}(\mathbf{x}_i) = \sum_{j=1}^{n} \alpha_j \mathbf{T}(\mathbf{y}_j) = \sum_{j=1}^{n} \alpha_j \mathbf{x}_j,$$

which means $[\mathbf{x}_i]_{\mathcal{B}'} = [\mathbf{T}(\mathbf{x}_i)]_{\mathcal{B}}$, so,

$$[\mathbf{T}]_{\mathcal{B}} = \Big([\mathbf{T}(\mathbf{x}_1)]_{\mathcal{B}} \; [\mathbf{T}(\mathbf{x}_2)]_{\mathcal{B}} \; \cdots [\mathbf{T}(\mathbf{x}_n)]_{\mathcal{B}} \Big) = \Big([\mathbf{x}_1]_{\mathcal{B}'} \; [\mathbf{x}_2]_{\mathcal{B}'} \cdots [\mathbf{x}_n]_{\mathcal{B}'} \Big) = [\mathbf{T}]_{\mathcal{B}'}.$$

The fact that $[\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}}$ follows because $[\mathbf{I}(\mathbf{x}_i)]_{\mathcal{B}'} = [\mathbf{x}_i]_{\mathcal{B}'}$. The matrix

$$\mathbf{P} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'} = [\mathbf{T}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}'}$$

- $[\mathbf{v}]_{\mathcal{B}'} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'}[\mathbf{v}]_{\mathcal{B}} = \mathbf{P}[\mathbf{v}]_{\mathcal{B}}$.
- $\mathbf{P}$ is nonsingular because $\mathbf{T}$ is invertible($\mathbf{T}^{-1}(\mathbf{x}_i) = \mathbf{y}_i$).
- $[\mathbf{T}^{-1}]_{\mathcal{B}} = [\mathbf{T}]_{\mathcal{B}}^{-1} = \mathbf{P}^{-1}$.

- The change of basis operator $\mathbf{T}$ acts as

$$\mathbf{T}(new\ basis) = old\ basis.$$

- While the change of basis matrix $\mathbf{P}$ acts as

$$new\ coordinates = \mathbf{P}(old\ coordinates).$$

- **Example:** For the space $\mathcal{P}_2$ of polynomials of degree 2 or less, determine the change of basis matrix $\mathbf{P}$ form $\mathcal{B}$ to $\mathcal{B}'$, where

$$\mathcal{B} = \{1, t, t^2\} \quad and \quad \mathcal{B}' = \{1, 1+t, 1+t+t^2\},$$

and then find the coordinates of $q(t) = 3 + 2t + 4t^2$ relative to $\mathcal{B}'$.

- **Solution:** The change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$ is

$$\mathbf{P} = \left([x_1]_{\mathcal{B}'} \mid [x_2]_{\mathcal{B}'} \mid [x_3]_{\mathcal{B}'}\right).$$

In this case, $\mathbf{x}_1 = 1$, $\mathbf{x}_2 = t$, and $\mathbf{x}_3 = t^2$, and $\mathbf{y}_1 = 1$, $\mathbf{y}_2 = 1 + t$, and $\mathbf{y}_3 = 1 + t + t^2$, so the coordinates $[\mathbf{x}_i]_{\mathcal{B}'}$ are computed as follows:

$$\begin{aligned} 1 &= 1(1) + 0(1+t) + 0(1+t+t^2) = 1\mathbf{y}_1 + 0\mathbf{y}_2 + 0\mathbf{y}_3, \\ t &= -1(1) + 1(1+t) + 0(1+t+t^2) = -1\mathbf{y}_1 + 1\mathbf{y}_2 + 0\mathbf{y}_3, \\ t^2 &= 0(1) - 1(1+t) + 1(1+t+t^2) = 0\mathbf{y}_1 - 1\mathbf{y}_2 + 1\mathbf{y}_3. \end{aligned}$$

Therefore,

$$\mathbf{P} = \left([\mathbf{x}_1]_{\mathcal{B}'} \,\middle|\, [\mathbf{x}_2]_{\mathcal{B}'} \,\middle|\, [\mathbf{x}_3]_{\mathcal{B}'} \right) = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix},$$

and the coordinates of $\mathbf{q} = q(t) = 3 + 2t + 4t^2$ with respect to $\mathcal{B}'$ are

$$[\mathbf{q}]_{\mathcal{B}'} = \mathbf{P}[\mathbf{q}]_{\mathcal{B}} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \\ 4 \end{pmatrix}.$$

To independently check that these coordinates are correct, simply verify that

$$q(t) = 1(1) - 2(1+t) + 4(1+t+t^2).$$

- It's now rather easy to describe how the coordinate matrix of a linear operator changes as the underlying basis changes.

Changing Matrix Coordinates

Let $\mathbf{A}$ be a linear operator on $\mathcal{V}$, and let $\mathcal{B}$ and $\mathcal{B}'$ be two bases for $\mathcal{V}$. The coordinate matrices $[\mathbf{A}]_{\mathcal{B}}$ and $[\mathbf{A}]_{\mathcal{B}'}$ are related as follows.

$$[\mathbf{A}]_{\mathcal{B}} = \mathbf{P}^{-1}[\mathbf{A}]_{\mathcal{B}'}\mathbf{P}, \quad \text{where} \quad \mathbf{P} = [\mathbf{I}]_{\mathcal{B}\mathcal{B}'}$$

is the change of basis matrix from $\mathcal{B}$ to $\mathcal{B}'$. Equivalently,

$$[\mathbf{A}]_{\mathcal{B}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{B}}\mathbf{Q}, \quad \text{where} \quad \mathbf{Q} = [\mathbf{I}]_{\mathcal{B}'\mathcal{B}} = \mathbf{P}^{-1}$$

is the change of basis matrix from $\mathcal{B}'$ to $\mathcal{B}$.

Problem: Consider the linear operator $\mathbf{A}(x, y) = (y,\ -2x + 3y)$ on $\Re^2$ along with the two bases

$$\mathcal{S} = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} \quad \text{and} \quad \mathcal{S}' = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix} \right\}.$$

First compute the coordinate matrix $[\mathbf{A}]_{\mathcal{S}}$ as well as the change of basis matrix $\mathbf{Q}$ from $\mathcal{S}'$ to $\mathcal{S}$, and then use these two matrices to determine $[\mathbf{A}]_{\mathcal{S}'}$.

Solution: The matrix of $\mathbf{A}$ relative to $\mathcal{S}$ is obtained by computing

$$\begin{aligned} \mathbf{A}(\mathbf{e}_1) &= \mathbf{A}(1, 0) = (0,\ -2) = (0)\mathbf{e}_1 + (-2)\mathbf{e}_2, \\ \mathbf{A}(\mathbf{e}_2) &= \mathbf{A}(0, 1) = (1, 3) = (1)\mathbf{e}_1 + (3)\mathbf{e}_2, \end{aligned}$$

so that $[\mathbf{A}]_{\mathcal{S}} = \left([\mathbf{A}(\mathbf{e}_1)]_{\mathcal{S}} \,\middle|\, [\mathbf{A}(\mathbf{e}_2)]_{\mathcal{S}} \right) = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}$.

The change of basis matrix from $\mathcal{S}'$ to $\mathcal{S}$ is

$$\mathbf{Q} = \left([\mathbf{y}_1]_{\mathcal{S}} \,\middle|\, [\mathbf{y}_2]_{\mathcal{S}} \right) = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

and the matrix of $\mathbf{A}$ with respect to $\mathcal{S}'$ is

$$[\mathbf{A}]_{\mathcal{S}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{S}}\mathbf{Q} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}.$$

Similarity

- Matrices $\mathbf{B}_{n\times n}$ and $\mathbf{C}_{n\times n}$ are said to be ***similar matrices*** whenever there exists a nonsingular matrix $\mathbf{Q}$ such that $\mathbf{B}=\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}$. We write $\mathbf{B}\simeq\mathbf{C}$ to denote that $\mathbf{B}$ and $\mathbf{C}$ are similar.
- The linear operator $f:\Re^{n\times n}\to\Re^{n\times n}$ defined by $f(\mathbf{C})=\mathbf{Q}^{-1}\mathbf{C}\mathbf{Q}$ is called a ***similarity transformation.***

- Any two coordinate matrices of a given linear operator must be similar.
- Must any two similar matrices be coordinate matrices of the same linear operator?
- Similar matrices represent the same linear operator.
- So the coordinate-independent properties are exactly the ones that are **similarity invariant** (invariant under similarity transformations).
- Determining and studying similarity invariants is an important part of linear algebra and matrix theory.

Invariant Subspaces

For a linear operator $\mathbf{T}$ on a vector space $\mathcal{V}$, and for $\mathcal{X} \subseteq \mathcal{V}$,

$$\mathbf{T}(\mathcal{X}) = \{\mathbf{T}(\mathbf{x}) \mid \mathbf{x} \in \mathcal{X}\}$$

is the set of all possible images of vectors from $\mathcal{X}$ under the transformation $\mathbf{T}$. Notice that $\mathbf{T}(\mathcal{V}) = R(\mathbf{T})$. When $\mathcal{X}$ is a subspace of $\mathcal{V}$, it follows that $\mathbf{T}(\mathcal{X})$ is also a subspace of $\mathcal{V}$, but $\mathbf{T}(\mathcal{X})$ is usually not related to $\mathcal{X}$. However, in some special cases it can happen that $\mathbf{T}(\mathcal{X}) \subseteq \mathcal{X}$, and such subspaces are the focus of this section.

Invariant Subspaces

- For a linear operator $\mathbf{T}$ on $\mathcal{V}$, a subspace $\mathcal{X} \subseteq \mathcal{V}$ is said to be an ***invariant subspace*** under $\mathbf{T}$ whenever $\mathbf{T}(\mathcal{X}) \subseteq \mathcal{X}$.
- In such a situation, $\mathbf{T}$ can be considered as a linear operator on $\mathcal{X}$ by forgetting about everything else in $\mathcal{V}$ and restricting $\mathbf{T}$ to act only on vectors from $\mathcal{X}$. Hereafter, this ***restricted operator*** will be denoted by $\mathbf{T}_{/\mathcal{X}}$.

Problem: For

$$\mathbf{A} = \begin{pmatrix} 4 & 4 & 4 \\ -2 & -2 & -5 \\ 1 & 2 & 5 \end{pmatrix}, \qquad \mathbf{x}_1 = \begin{pmatrix} 2 \\ -1 \\ 0 \end{pmatrix}, \qquad \text{and} \qquad \mathbf{x}_2 = \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix},$$

show that the subspace $\mathcal{X}$ spanned by $\mathcal{B} = \{\mathbf{x}_1, \ \mathbf{x}_2\}$ is an invariant subspace under $\mathbf{A}$. Then describe the restriction $\mathbf{A}_{/\mathcal{X}}$ and determine the coordinate matrix of $\mathbf{A}_{/\mathcal{X}}$ relative to $\mathcal{B}$.

Solution: Observe that $\mathbf{A}\mathbf{x}_1 = 2\mathbf{x}_1 \in \mathcal{X}$ and $\mathbf{A}\mathbf{x}_2 = \mathbf{x}_1 + 2\mathbf{x}_2 \in \mathcal{X}$, so the image of any $\mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 \in \mathcal{X}$ is back in $\mathcal{X}$ because

$$\mathbf{A}\mathbf{x} = \mathbf{A}(\alpha\mathbf{x}_1 + \beta\mathbf{x}_2) = \alpha\mathbf{A}\mathbf{x}_1 + \beta\mathbf{A}\mathbf{x}_2 = 2\alpha\mathbf{x}_1 + \beta(\mathbf{x}_1 + 2\mathbf{x}_2) = (2\alpha + \beta)\mathbf{x}_1 + 2\beta\mathbf{x}_2.$$

This equation completely describes the action of $\mathbf{A}$ restricted to $\mathcal{X}$, so

$$\mathbf{A}_{/\mathcal{X}}(\mathbf{x}) = (2\alpha + \beta)\mathbf{x}_1 + 2\beta\mathbf{x}_2 \quad \text{for each} \quad \mathbf{x} = \alpha\mathbf{x}_1 + \beta\mathbf{x}_2 \in \mathcal{X}.$$

Since $\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_1) = 2\mathbf{x}_1$ and $\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_2) = \mathbf{x}_1 + 2\mathbf{x}_2$, we have

$$\left[\mathbf{A}_{/\mathcal{X}}\right]_{\mathcal{B}} = \left(\left[\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_1)\right]_{\mathcal{B}} \,\middle|\, \left[\mathbf{A}_{/\mathcal{X}}(\mathbf{x}_2)\right]_{\mathcal{B}} \right) = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}.$$

The invariant subspaces for a linear operator $\mathbf{T}$ are important because they produce simplified coordinate matrix representations of $\mathbf{T}$. To understand how this occurs, suppose $\mathcal{X}$ is an invariant subspace under $\mathbf{T}$, and let

$$\mathcal{B}_{\mathcal{X}} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r\}$$

be a basis for $\mathcal{X}$ that is part of a basis

$$\mathcal{B} = \{\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_r, \mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_q\}$$

for the entire space $\mathcal{V}$. To compute $[\mathbf{T}]_{\mathcal{B}}$, recall from the definition of coordinate matrices that

$$[\mathbf{T}]_{\mathcal{B}} = \Big([\mathbf{T}(\mathbf{x}_1)]_{\mathcal{B}} \Big| \cdots \Big| [\mathbf{T}(\mathbf{x}_r)]_{\mathcal{B}} \Big| [\mathbf{T}(\mathbf{y}_1)]_{\mathcal{B}} \Big| \cdots \Big| [\mathbf{T}(\mathbf{y}_q)]_{\mathcal{B}} \Big).$$

Because each $\mathbf{T}(\mathbf{x}_j)$ is contained in $\mathcal{X}$, only the first r vectors from $\mathcal{B}$ are needed to represent each $\mathbf{T}(\mathbf{x}_j)$, so, for $j = 1, 2, \ldots, r$,

$$\mathbf{T}(\mathbf{x}_j) = \sum_{i=1}^{r} \alpha_{ij}\mathbf{x}_i \quad \text{and} \quad [\mathbf{T}(\mathbf{x}_j)]_{\mathcal{B}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{rj} \\ 0 \\ \vdots \\ 0 \end{pmatrix}.$$

The space

$$\mathcal{Y} = span\{\mathbf{y}_1, \mathbf{y}_2, \ldots, \mathbf{y}_q\}$$

may not be an invariant subspace for $\mathbf{T}$, so all the basis vectors in $\mathcal{B}$ may be needed to represent the $\mathbf{T}(\mathbf{y}_j)$'s. Consequently, for $j = 1, 2, \ldots, q$,

$$\mathbf{T}(\mathbf{y}_j) = \sum_{i=1}^{r} \beta_{ij}\mathbf{x}_i + \sum_{i=1}^{q} \gamma_{ij}\mathbf{y}_i \quad \text{and} \quad [\mathbf{T}(\mathbf{y}_j)]_{\mathcal{B}} = \begin{pmatrix} \beta_{1j} \\ \vdots \\ \beta_{rj} \\ \gamma_{1j} \\ \vdots \\ \gamma_{qj} \end{pmatrix}.$$

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \alpha_{11} & \cdots & \alpha_{1r} & \beta_{11} & \cdots & \beta_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \cdots & \alpha_{rr} & \beta_{r1} & \cdots & \beta_{rq} \\ 0 & \cdots & 0 & \gamma_{11} & \cdots & \gamma_{1q} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & \gamma_{q1} & \cdots & \gamma_{qq} \end{pmatrix}.$$

- The equation $\mathbf{T}(\mathbf{x}_j) = \sum_{i=1}^{r} \alpha_{ij}\mathbf{x}_i$ mean that

$$[\mathbf{T}_{/\mathcal{X}}]_{\mathcal{B}_{\mathcal{X}}} = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1r} \\ \alpha_{21} & \alpha_{22} & \cdots & \alpha_{2r} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_{r1} & \alpha_{r2} & \cdots & \alpha_{rr} \end{pmatrix}$$

- So, we have

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} [\mathbf{T}_{/\mathcal{X}}]_{\mathcal{B}_{\mathcal{X}}} & \mathbf{B}_{r\times q} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}.$$

- This says that the matrix representation for $\mathbf{T}$ can be made to be block triangular whenever a basis for an invariant subspace is available.
- The more invariant subspaces we can find, the more tools we have to construct simplified matrix representations.

This notion easily generalizes in the sense that if $\mathcal{B} = \mathcal{B}_{\mathcal{X}} \cup \mathcal{B}_{\mathcal{Y}} \cup \cdots \cup \mathcal{B}_{\mathcal{Z}}$ is a basis for $\mathcal{V}$, where $\mathcal{B}_{\mathcal{X}}, \mathcal{B}_{\mathcal{Y}}, \ldots, \mathcal{B}_{\mathcal{Z}}$ are bases for invariant subspaces under $\mathbf{T}$ that have dimensions $r_1, r_2, \ldots, r_k$, respectively, then $[\mathbf{T}]_{\mathcal{B}}$ has the block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r_1 \times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2 \times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k \times r_k} \end{pmatrix},$$

where

$$\mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_x}, \quad \mathbf{B} = \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\mathcal{B}_y}, \quad \ldots, \quad \mathbf{C} = \left[\mathbf{T}_{/\mathcal{Z}}\right]_{\mathcal{B}_z}.$$

The situations discussed above are also reversible in the sense that if the matrix representation of $\mathbf{T}$ has a block-triangular form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r \times r} & \mathbf{B}_{r \times q} \\ \mathbf{0} & \mathbf{C}_{q \times q} \end{pmatrix}$$

relative to some basis

$$\mathcal{B} = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r, \mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\},$$

then the r-dimensional subspace $\mathcal{U} = span\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\}$ spanned by the first r vectors in $\mathcal{B}$ must be an invariant subspace under $\mathbf{T}$. Furthermore, if the matrix representation of $\mathbf{T}$ has a block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}$$

relative to $\mathcal{B}$, then both

$$\mathcal{U} = span\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_r\} \quad \text{and} \quad \mathcal{W} = span\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_q\}$$

must be invariant subspaces for $\mathbf{T}$.

The general statement concerning invariant subspaces and coordinate matrix representations is given below.

Invariant Subspaces and Matrix Representations

Let $\mathbf{T}$ be a linear operator on an n-dimensional space $\mathcal{V}$, and let $\mathcal{X}, \mathcal{Y}, \ldots, \mathcal{Z}$ be subspaces of $\mathcal{V}$ with respective dimensions $r_1, r_2, \ldots, r_k$ and bases $\mathcal{B}_{\mathcal{X}}, \mathcal{B}_{\mathcal{Y}}, \ldots, \mathcal{B}_{\mathcal{Z}}$. Furthermore, suppose that $\sum_i r_i = n$ and $\mathcal{B} = \mathcal{B}_{\mathcal{X}} \cup \mathcal{B}_{\mathcal{Y}} \cup \cdots \cup \mathcal{B}_{\mathcal{Z}}$ is a basis for $\mathcal{V}$.

- The subspace $\mathcal{X}$ is an invariant subspace under $\mathbf{T}$ if and only if $[\mathbf{T}]_{\mathcal{B}}$ has the block-triangular form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r_1\times r_1} & \mathbf{B} \\ \mathbf{0} & \mathbf{C} \end{pmatrix}, \quad \text{in which case} \quad \mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_{\mathcal{X}}}.$$

- The subspaces $\mathcal{X}, \mathcal{Y}, \ldots, \mathcal{Z}$ are all invariant under $\mathbf{T}$ if and only if $[\mathbf{T}]_{\mathcal{B}}$ has the block-diagonal form

$$[\mathbf{T}]_{\mathcal{B}} = \begin{pmatrix} \mathbf{A}_{r_1 \times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2 \times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k \times r_k} \end{pmatrix},$$

in which case

$$\mathbf{A} = \left[\mathbf{T}_{/\mathcal{X}}\right]_{\mathcal{B}_x}, \quad \mathbf{B} = \left[\mathbf{T}_{/\mathcal{Y}}\right]_{\mathcal{B}_y}, \quad \ldots, \quad \mathbf{C} = \left[\mathbf{T}_{/\mathcal{Z}}\right]_{\mathcal{B}_z}.$$

An important corollary concerns the special case in which the linear operator $\mathbf{T}$ is in fact an $n \times n$ matrix and $\mathbf{T}(\mathbf{v}) = \mathbf{T}\mathbf{v}$ is a matrix–vector multiplication.

Triangular and Diagonal Block Forms

When $\mathbf{T}$ is an $n \times n$ matrix, the following two statements are true.

- $\mathbf{Q}$ is a nonsingular matrix such that

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \begin{pmatrix} \mathbf{A}_{r\times r} & \mathbf{B}_{r\times q} \\ \mathbf{0} & \mathbf{C}_{q\times q} \end{pmatrix}$$

  if and only if the first r columns in $\mathbf{Q}$ span an invariant subspace under $\mathbf{T}$.

- $\mathbf{Q}$ is a nonsingular matrix such that

$$\mathbf{Q}^{-1}\mathbf{T}\mathbf{Q} = \begin{pmatrix} \mathbf{A}_{r_1\times r_1} & \mathbf{0} & \cdots & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_{r_2\times r_2} & \cdots & \mathbf{0} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{0} & \mathbf{0} & \cdots & \mathbf{C}_{r_k\times r_k} \end{pmatrix}$$

  if and only if $\mathbf{Q} = \left(\mathbf{Q}_1 \,|\, \mathbf{Q}_2 \,|\, \cdots \,|\, \mathbf{Q}_k\right)$ in which $\mathbf{Q}_i$ is $n \times r_i$, and the columns of each $\mathbf{Q}_i$ span an invariant subspace under $\mathbf{T}$.

Problem: Find all subspaces of $\Re^2$ that are invariant under

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ -2 & 3 \end{pmatrix}.$$

Solution: The trivial subspace $\{\mathbf{0}\}$ is the only zero-dimensional invariant subspace, and the entire space $\Re^2$ is the only two-dimensional invariant subspace. The real problem is to find all one-dimensional invariant subspaces. If $\mathcal{M}$ is a one-dimensional subspace spanned by $\mathbf{x} \neq \mathbf{0}$ such that $\mathbf{A}(\mathcal{M}) \subseteq \mathcal{M}$, then

$$\mathbf{A}\mathbf{x} \in \mathcal{M} \implies \text{there is a scalar } \lambda \text{ such that } \mathbf{A}\mathbf{x} = \lambda\mathbf{x} \implies (\mathbf{A} - \lambda\mathbf{I})\,\mathbf{x} = \mathbf{0}.$$

In other words, $\mathcal{M} \subseteq N(\mathbf{A} - \lambda\mathbf{I})$. Since $\dim \mathcal{M} = 1$, it must be the case that $N(\mathbf{A} - \lambda\mathbf{I}) \neq \{\mathbf{0}\}$, and consequently λ must be a scalar such that $(\mathbf{A} - \lambda\mathbf{I})$ is a singular matrix. Row operations produce

$$\mathbf{A} - \lambda\mathbf{I} = \begin{pmatrix} -\lambda & 1 \\ -2 & 3-\lambda \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 3-\lambda \\ -\lambda & 1 \end{pmatrix} \longrightarrow \begin{pmatrix} -2 & 3-\lambda \\ 0 & 1+(\lambda^2-3\lambda)/2 \end{pmatrix},$$

and it is clear that $(\mathbf{A} - \lambda\mathbf{I})$ is singular if and only if $1 + (\lambda^2 - 3\lambda)/2 = 0$—i.e., if and only if λ is a root of

$$\lambda^2 - 3\lambda + 2 = 0.$$

Thus $\lambda = 1$ and $\lambda = 2$, and straightforward computation yields the two one-dimensional invariant subspaces

$$\mathcal{M}_1 = N(\mathbf{A} - \mathbf{I}) = span\left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}\right\} \quad \text{and} \quad \mathcal{M}_2 = N(\mathbf{A} - \mathbf{2I}) = span\left\{\begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}.$$

In passing, notice that $\mathcal{B} = \left\{\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$ is a basis for $\Re^2$, and

$$[\mathbf{A}]_{\mathcal{B}} = \mathbf{Q}^{-1}\mathbf{A}\mathbf{Q} = \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \quad \text{where} \quad \mathbf{Q} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

In general, scalars λ for which $(\mathbf{A} - \lambda\mathbf{I})$ is singular are called the ***eigenvalues*** of $\mathbf{A}$, and the nonzero vectors in $N(\mathbf{A} - \lambda\mathbf{I})$ are known as the associated ***eigenvectors*** for $\mathbf{A}$. As this example indicates, eigenvalues and eigenvectors are of fundamental importance in identifying invariant subspaces and reducing matrices by means of similarity transformations.

Exercise

1. Determine which of the following functions are linear operators on $\mathfrak{R}^2$
   (a) $\mathbf{T}(x,y) = (x, 1+y)$,
   (b) $\mathbf{T}(x,y) = (0, xy)$,
   (c) $\mathbf{T}(x,y) = (x^2, y^2)$.
2. Explain why $\mathbf{T}(0) = 0$ for every linear transformation $\mathbf{T}$.
3. Explain why rank is a similarity invariant.
4. $\mathbf{A}(x,y,z) = (x+2y-z, -y, x+7z)$ is a linear operator on $\mathfrak{R}^3$.
   (a) Determine $[\mathbf{A}]_{\mathcal{S}}$, where $\mathcal{S}$ is the standard basis.
   (b) Determine $[\mathbf{A}]_{\mathcal{S}'}$ as well as the nonsingular matrix $\mathbf{Q}$ such that

$$[\mathbf{A}]_{\mathcal{S}'} = \mathbf{Q}^{-1}[\mathbf{A}]_{\mathcal{S}}\mathbf{Q} \quad for \quad \mathcal{S}' = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}.$$

5. Let $\mathbf{T}$ be an arbitrary linear operator on a vector space $\mathcal{V}$.
   (a) Is the trivial subspace $\{\mathbf{o}\}$ invariant under $\mathbf{T}$?
   (b) Is the entire space $\mathcal{V}$ invariant under $\mathbf{T}$?

6. Describe all of the subspace that are invariant under the identity operator $\mathbf{I}$ on a space $\mathcal{V}$.
7. Let $\mathbf{T}$ be the linear operator on $\mathfrak{R}^4$ defined by

$$\mathbf{T}(x_1, x_2, x_3, x_4) = (x_1 + x_2 + 2x_3 - x_4,\ x_2 + x_4,\ 2x_3 - x_4,\ x_3 + x_4),$$

and let $\mathcal{X} = span\{e_1, e_2\}$ be the subspace that is spanned by the first two unit vectors in $\mathfrak{R}^4$.

(a) Explain why $\mathcal{X}$ is invariant under $\mathbf{T}$.

(b) Determine $\left[\mathbf{T}_{/\mathcal{X}}\right]_{\{e_1, e_2\}}$.

(c) Describe the structure of $[\mathbf{T}]_{\mathcal{B}}$, where $\mathcal{B}$ is any basis obtained from an extension of $\{e_1, e_2\}$.