Intro to Linear Algebra MAT 5230

Wm C Bauldry

Autumn Semester, 2005

Table of Contents

Table of Contents

Table of Contents

1. Algebraic Structures

Definition 1 *A* Group *is a pair* $\{X; \cdot\}$ *such that*

1. *"$\cdot$" is* closed *on* X*.*
2. *"$\cdot$" is* associative *on* X*.*
3. *There is an identity* $e \in X$ *(w.r.t. "$\cdot$").*
4. *Every element* $a \in X$ *has an inverse* a^{-1} *(w.r.t. "$\cdot$").*

Definition 2 *A* Ring *is a triple* $\{X; +, \cdot\}$ *such that*

1. $\{X; +\}$ *is an* Abelian group*.*
2. $\{X; \cdot\}$ *is a* semigroup *(lacks identity and inverses).*
3. *"$\cdot$" distributes over "$+$".*

Algebraic Structures

Definition 3 *A* Field *is a triple* $\{X;+,\cdot\}$ *such that*

1. $\{X;+,\cdot\}$ *is a* ring*.*
2. $\{X^{\#};\cdot\}$ *is an* Abelian group *where* $X^{\#} = X - \{0\}$*.*

Definition 4 *A* Vector Space *is an Abelian group* $\{X;+\}$ *over a field* $\{F;+,\cdot\}$ *with a* scalar product $F \times X \to X$*. For* $\alpha, \beta \in F$ *and* $x, y \in X$*,*

1. $\alpha(x+y) = \alpha x + \alpha y$
2. $(\alpha+\beta)x = \alpha x + \beta x$
3. $(\alpha\beta)x = \alpha(\beta x)$
4. $1x = x$

Field

Definition 3 (Field) *Let* $F \neq \emptyset$ *be a set with addition* "$+$"*:*$X \times X \to X$ *and multiplication* "$\cdot$"*:*$F \times X \to X$*. Then* $\{F; +, \cdot\}$ *with the operations forms a* field *if the following axioms are satisfied:*

1. $x + y = y + x,\ x \cdot y = y \cdot x$ — commutative laws
2. $x + (y + z) = (x + y) + z,\ x \cdot (y \cdot z) = (x \cdot y) \cdot z$ — associative laws
3. $\exists$ *unique element* 0 *satisfying* $0 + x = x$ — additive identity
4. *To each* x*,* $\exists$ *a unique* $-x$ *so that* $x + (-x) = 0$ — additive inverse
5. *There is a unique element* 1 *satisfying* $1 \cdot x = x$ — mult. identity
6. *To each* $x \neq 0$*,* $\exists$ *a unique* x^{-1} *so that* $x \cdot x^{-1} = 1$ — mult. inverse
7. $x \cdot (y + z) = x \cdot y + x \cdot z$ — "$\cdot$" over "$+$" distributive law

Examples of Fields

1. $\mathbb{Q}$, $\mathbb{R}$, and $\mathbb{C}$ are fields.
2. $\mathbb{Z}$ is *not* a field. (*Why?*)
3. Let p be a prime. Then $\mathbb{Z}_p$ is a p-element field.
4. $\mathbb{Q}\left[\sqrt{2}\right] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Q}\}$ is a field.
5. $\mathbb{Z}\left[\sqrt{2}\right] = \{a + b\sqrt{2} \mid a, b \in \mathbb{Z}\}$ is *not* a field. (*Why?*)
6. $\mathbb{Q}\left[\sqrt[3]{3}\right] = \{a + b\sqrt[3]{3} + c\sqrt[3]{3^2} \mid a, b, c \in \mathbb{Q}\}$ is a field.
7. $\mathbb{Z}_p[i]$, p is prime, is a field (with p^2 elements).

Vector Space

Definition 4 (Vector Space) *Let* $X \neq \emptyset$ *be a set (*vectors*) and* F *be a field (*scalars*) with* vector addition *"*$+$*":*$X \times X \to X$ *and* scalar multiplication *"*$\cdot$*":*$F \times X \to X$*. Then* X *and* F *with the operations forms a* vector space *(or* linear space*), "*X *is a* vector space *over* F*," if the following axioms are satisfied:*

1. $x + y = y + x$ — commutative law
2. $x + (y + z) = (x + y) + z$ — associative law
3. *There is a unique vector* 0 *satisfying* $0 + x = x$ — 'zero vector,' identity
4. $\alpha(x + y) = \alpha x + \alpha y$ — scalar "$\cdot$" over vector "$+$" distributive law
5. $(\alpha + \beta)x = \alpha x + \beta x$ — scalar "$+$" over scalar "$\cdot$" distributive law
6. $(\alpha\beta)x = \alpha(\beta x)$ — scalar homogeneity
7. $0x = 0$ — scalar-vector additive identity relation (*implied by 5.*)
8. $1x = x$ — scalar-vector multiplicative identity relation

Examples of Vector Spaces

1. Let $n \in \mathbb{Z}^+$. Then $\mathbb{Q}^n$, $\mathbb{R}^n$, and $\mathbb{C}^n$ are vector spaces.
2. Let $n \in \mathbb{Z}^+$. Then $\mathbb{P}^n$, the polynomials (real or complex) of degree less than or equal to n, forms a vector space.
3. $\mathbb{Z}_2 \times \mathbb{Z}_2 \times \mathbb{Z}_2$ is a vector space.
4. Let F be a field and $n \in \mathbb{Z}^+$. Then F^n is a vector space.
5. Let $M_{m\times n}$ be the $m \times n$ matrices with entries in a field F with componentwise addition and scalar multiplication.
6. Let $K \subseteq \mathbb{R}$ be a closed interval. Then $C(K)$, the continuous real-valued functions on K form a vector space.
7. Let $O \subseteq \mathbb{R}$ be an open interval. Then $C^1(O)$, the continuously differentiable real-valued functions on O, form a vector space.

Homomorphisms

Definition 5 (Group Homomorphism) *Let* $\{X; +_X\}$ *and* $\{Y; +_Y\}$ *be two groups with* $\rho : X \to Y$. *Then* ρ *is a* homomorpism *iff*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2)$$

Definition 6 (Ring Homomorphism) *Let* $\{X; +_X, \cdot_X\}$ *and* $\{Y; +_Y, \cdot_Y\}$ *be two rings with* $\rho : X \to Y$. *Then* ρ *is a* homomorpism *iff*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2)$$

$$\rho(x_1 \cdot_X x_2) = \rho(x_1) \cdot_Y \rho(x_2)$$

Vector Space Homomorphism

Definition 7 (Linear Transformation) *Let* X *and* Y *be vector spaces over the same field* F*. Then the relation* $\rho : X \to Y$ *is a* linear transformation *if and only if for every* $\alpha \in F$ *and* $x_1, x_2 \in X$*, it follows that:*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2) \tag{1}$$

$$\rho(\alpha \cdot x_1) = \alpha \cdot \rho(x_1) \tag{2}$$

Linear Transformation

$$
\begin{array}{ccc}
[x_1, x_2] & \xrightarrow{+} & x_1 + x_2 \\
\rho \big\downarrow & & \rho \big\downarrow \\
[\rho(x_1), \rho(x_2)] & \xrightarrow{+} & \begin{array}{c} \rho(x_1 + x_2) = \\ \rho(x_1) + \rho(x_2) \end{array}
\end{array}
\tag{1}
$$

$$
\begin{array}{ccc}
[\alpha, x_1] & \xrightarrow{\cdot} & \alpha \cdot x_1 \\
\rho \big\downarrow & & \rho \big\downarrow \\
[\alpha, \rho(x_1)] & \xrightarrow{\cdot} & \begin{array}{c} \rho(\alpha \cdot x_1) = \\ \alpha \cdot \rho(x_1) \end{array}
\end{array}
\tag{2}
$$

(Go to TOC)

2. Properties of Finite Fields

Theorem 1 *$\mathbb{Z}_p$ is a field if and only if p is prime.*

Theorem 2 *Let p be a prime and $n \in \mathbb{Z}^+$. Then there exists a finite field F with p^n elements.*

Theorem 3 *For any prime p and $n \in \mathbb{Z}^+$, there is (essentially) only one field with p^n elements.*
(The splitting field *of $x^{p^n} - x$ over the field $\mathbb{Z}_p$.)*

References

- *Elementary Modern Algebra*, R Thompson, Scott, Foresman, & Co.
- *Modern Algebra: A First Course*, H Hollister, Harper & Row.
- *Introduction to Modern Algebra*, H McCoy, Allyn and Bacon.
- *Modern Algebra*, F Ayers, Schaum's Outline Series, McGraw-Hill.
- *Basic Algebra I*, N Jacobson, Freeman.
- http://mathworld.wolfram.com/FiniteField.html
- http://www-math.cudenver.edu/~wcherowi/courses/finflds.html

(Go to TOC)

4. Properties of Vector Spaces

Theorem 4 *Let X be a vector space over the field F. Let $x, y, z \in X$ and $\alpha, \beta \in F$. Then*

1. *if $\alpha x = \alpha y$ and $\alpha \neq 0$, then $x = y$;*
2. *if $\alpha x = \beta x$ and $x \neq 0$, then $\alpha = \beta$;*
3. *if $x + y = x + z$, then $y = z$;*
4. *$\alpha \cdot 0 = 0$;*
5. *$\alpha(x - y) = \alpha x - \alpha y$ where $-y \stackrel{\Delta}{=} (-1) \cdot y$;*
6. *$(\alpha - \beta)x = \alpha x - \beta x$;*
7. *$x + y = 0$ implies that $x = -y$.*

More Examples of Vector Spaces

Sequence Vector Spaces

- $\mathbb{R}^\infty$ and $\mathbb{C}^\infty$
- Finitely non-zero real (or complex) sequences
- Null real (or complex) sequences
- Bounded real (or complex) sequences
- Convergent real (or complex) sequences

Function Vector Spaces

- $\mathbb{P} = \{$polynomials with real (or complex) coefficients$\}$
- $C([a,b]) = \{f \mid f : [a,b] \to \mathbb{R}$ is continuous$\}$ over $\mathbb{R}$
- $L_1([a,b]) = \{f \mid \int_a^b |f(t)|\, dt < \infty\}$ over $\mathbb{R}$

(Go to TOC)

5. Homomorphisms

Definition 5 (Group Homomorphism) *Let* $\{X; +_X\}$ *and* $\{Y; +_Y\}$ *be two groups with* $\rho : X \to Y$*. Then* ρ *is a* homomorpism *iff*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2)$$

Definition 6 (Ring Homomorphism) *Let* $\{X; +_X, \cdot_X\}$ *and* $\{Y; +_Y, \cdot_Y\}$ *be two rings with* $\rho : X \to Y$*. Then* ρ *is a* homomorpism *iff*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2)$$

$$\rho(x_1 \cdot_X x_2) = \rho(x_1) \cdot_Y \rho(x_2)$$

Vector Space Homomorphism

Definition 7 (Linear Transformation) *Let X and Y be vector spaces over the same field F. Then the relation $\rho : X \to Y$ is a* linear transformation *if and only if for every $\alpha \in F$ and $x_1, x_2 \in X$, it follows that:*

$$\rho(x_1 +_X x_2) = \rho(x_1) +_Y \rho(x_2) \tag{3}$$

$$\rho(\alpha \cdot x_1) = \alpha \cdot \rho(x_1) \tag{4}$$

Examples

1. Set $\phi : \mathbb{R}^2 \to \mathbb{R}^4$ by $\phi(x, y) = (x, 0, 0, y)$.
2. Set $\psi : \mathbb{R}^2 \to \mathbb{C}$ by $\psi(x, y) = x + i\, y$.

Linear Transformation

$$\begin{array}{ccc} [x_1, x_2] & \xrightarrow{+} & x_1 + x_2 \\ \rho \downarrow & & \rho \downarrow \\ [\rho(x_1), \rho(x_2)] & \xrightarrow{+} & \begin{array}{c} \rho(x_1 + x_2) = \\ \rho(x_1) + \rho(x_2) \end{array} \end{array} \tag{1}$$

$$\begin{array}{ccc} [\alpha, x_1] & \xrightarrow{\cdot} & \alpha \cdot x_1 \\ \rho \downarrow & & \rho \downarrow \\ [\alpha, \rho(x_1)] & \xrightarrow{\cdot} & \begin{array}{c} \rho(\alpha \cdot x_1) = \\ \alpha \cdot \rho(x_1) \end{array} \end{array} \tag{2}$$

(Go to TOC)

6. Subspace of a Vector Space

Definition 8 (Subspace) *Let X be a vector space over F and let $\emptyset \neq V \subseteq X$. Then V is a* subspace *of X iff*

1. *$\forall u, v \in V$, we have $u + v \in V$ (closed under addition)*
2. *$\forall \alpha \in F$, $\forall u \in V$, we have $\alpha u \in V$ (closed under scalar mult.)*

Theorem 5 *A subspace of a vector space is itself a vector space.*

Proof. Let V be a subspace of X. V is closed under vector addition and scalar multiplication by definition. All remaining vector space properties — with the exception of $0 \in V$ — are inherited from X. Let $v \in V$ (because $V \neq \emptyset$). Since $0 \in F$, then $0v = 0 \in V$. Thus V is a vector space. □

Note. Every vector space has at least 2 subspaces. What are they?

Examples of Subspaces

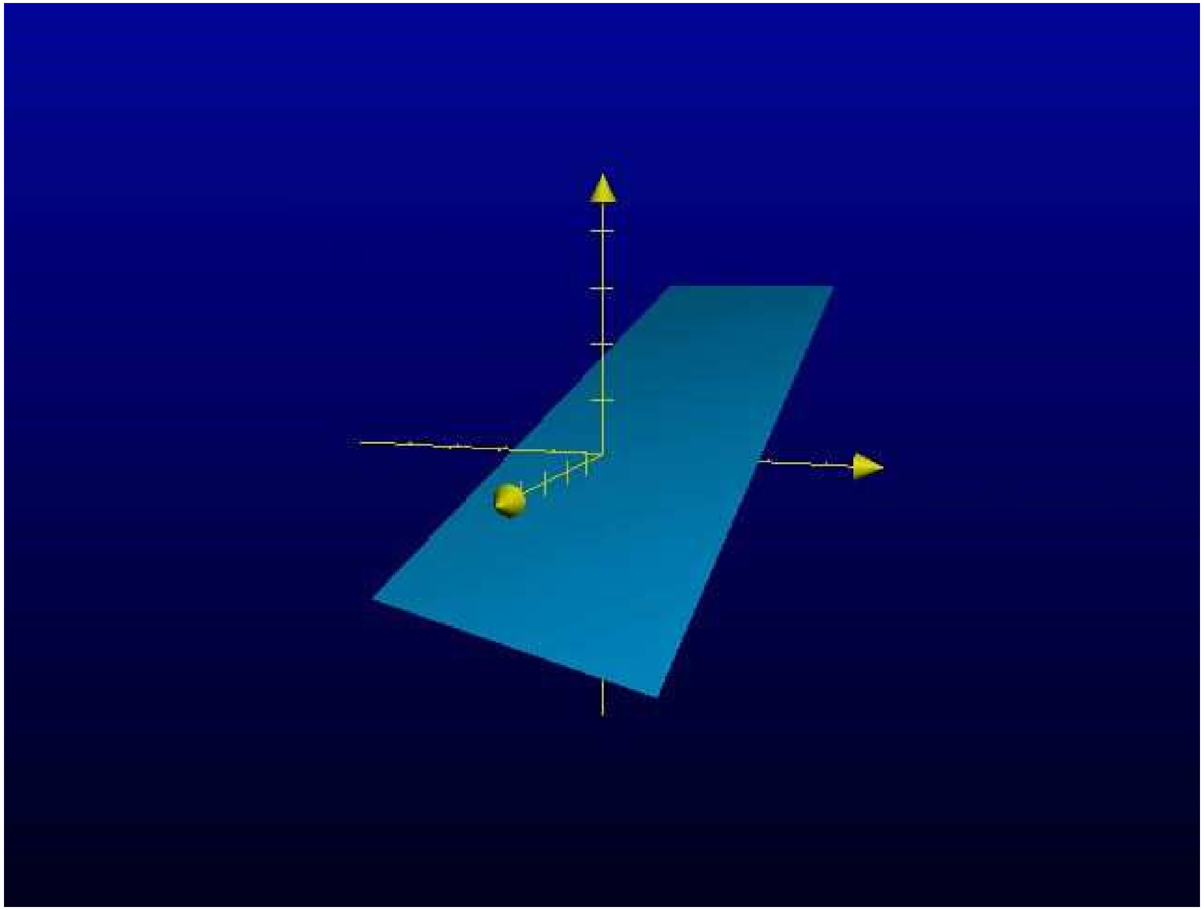
- $\{0\}$ and X are always subspaces of X
- $\mathbb{R}^2$ is a subspace[a] of $\mathbb{R}^3$, $\mathbb{C}^2$ is a subspace of $\mathbb{C}^3$.
- For $m < n$, we have that $\mathbb{R}^m$ is a subspace of $\mathbb{R}^n$
- For $m < n$, we have that $\mathbb{P}^m$ is a subspace of $\mathbb{P}^n$
- Is
  - $V_1 = \{(x, 1) \,|\, x, y \in \mathbb{R}\}$ a subspace of $\mathbb{R}^2$?
  - $V_2 = \{(x, y, x + y, 0) \,|\, x, y \in \mathbb{R}\}$ a subspace of $\mathbb{R}^4$?
  - $V_3 = \{(x, y, x + y + 2, 0) \,|\, x, y \in \mathbb{R}\}$ a subspace of $\mathbb{R}^4$?

---

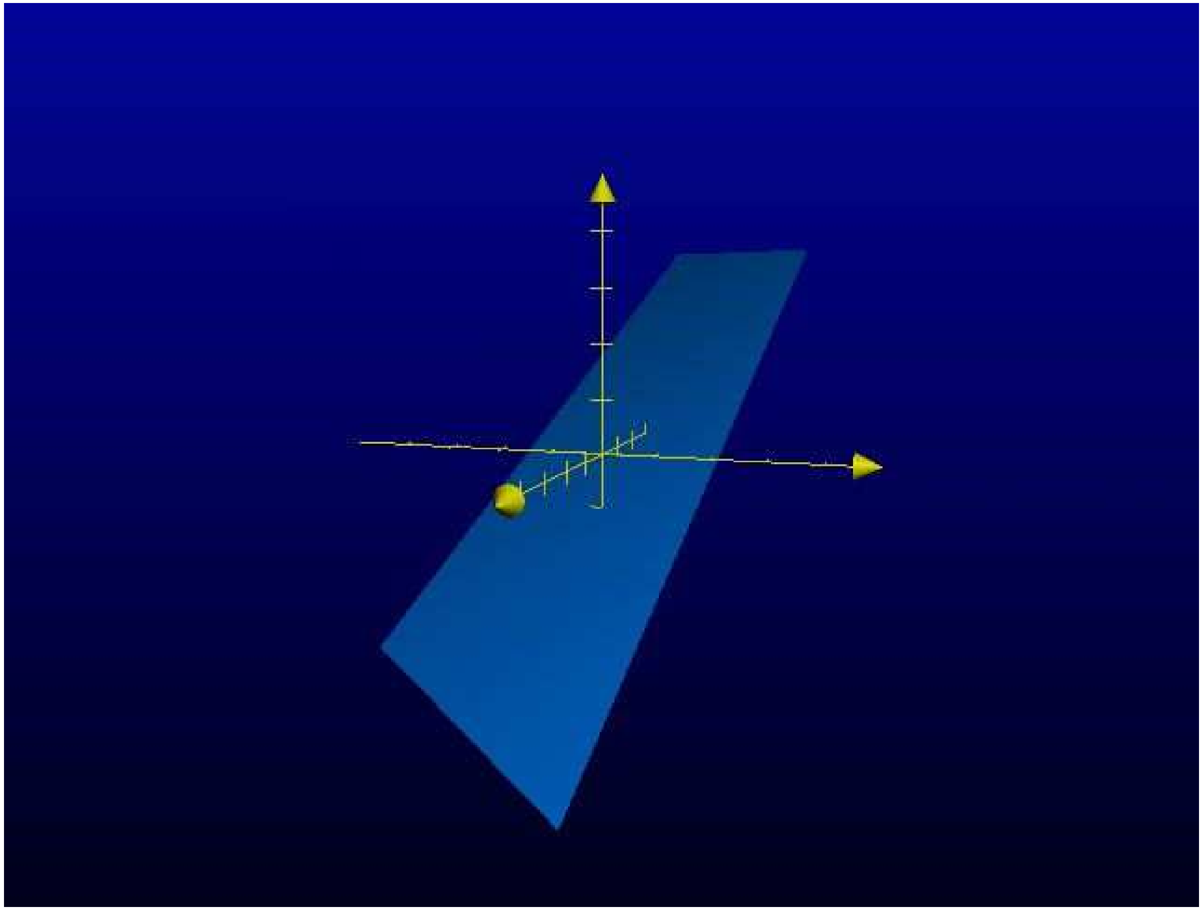
[a] Thinking of $\mathbb{R}^2$ as a subset such as $\{(x, y, 0) \,|\, x, y \in \mathbb{R}\}$, &c., of $\mathbb{R}^3$. Formally, $\mathbb{R}^2$ is isomorphic to a subspace of $\mathbb{R}^3$.

(Go to TOC)

Is This a Subspace?



Is This a Subspace?



(Go to TOC)

7. Operations with Subspaces

Theorem 6 *Let X be a vector space over F and let V_1 and V_2 be subspaces of X. Then $V = V_1 \cap V_2$ is a subspace.*

Pf. (Exercise.)

Theorem 7 *Let X be a vector space over F and let X_i for $i \in I$ be subspaces of X where I is some index set. Then $V = \bigcap_{i \in I} X_i$ is a subspace.*

Pf. (Easy closure calculations.)

NB: Unions (usually) or complements of subspaces do not form new subspaces.

Direct Sum

Definition 9 (Direct Sum) *Let* $X_1, X_2, \ldots, X_r$ *be subspaces of* X*. The set* $X_1 + X_2 + \cdots + X_r$ *forms the* direct sum $X_1 \oplus X_2 \oplus \cdots \oplus X_r$ *iff for every* x *in the sum, there is a unique set of* $x_i \in X_i$ *such that* $x = \sum_{i=1}^{r} x_i$*.*

Theorem 8 $X_1 + X_2 = X_1 \oplus X_2$ *if and only if* $X_1 \cap X_2 = \{0\}$*.*

Pf. Based on: Let $0 \neq v \in X_1 \cap X_2$. Then $v = v + 0 = 0 + v$ is two different ways to write v.

Note. $X_1 + X_2$ is a subspace; $X_1 \oplus X_2$ is a subspace that 'looks like' a direct product.

Subspaces of $\mathbb{R}^2$ and $\mathbb{R}^3$

Example 1 *Set* $X = \mathbb{R}^2$*. Let* X_1 *be given by the line* $y = x$ *and* X_2 *by the line* $y = -x$*. Then*

$$\{0\} = \underset{\text{subsp}}{X_1 \cap X_2} \subseteq \underset{\neg\text{subsp}}{X_1 \cup X_2} \subseteq \underset{\text{subsp}}{X_1 + X_2} = X_1 \oplus X_2 = \mathbb{R}^2$$

Example 2 *Set* $X = \mathbb{R}^3$*. The subspaces of* $\mathbb{R}^3$ *are:*

- $\{0\}$
- *A line* L *through the origin.*
- *The direct sum of two distinct lines through the origin* $L_1 \oplus L_2$ *yields a plane.*
- *The direct sum of three distinct non-coplanar lines through the origin* $L_1 \oplus L_2 \oplus L_3$ *yields* $\mathbb{R}^3$*.*

(Go to TOC)

8. Linear Combinations

Note: From now on, α_i, etc., will be elements of the base field F and x_i, y_i, etc., will be vectors from the space X.

Definition 10 (Finite Linear Combination) *Let $Y \subseteq X$. A vector $x \in X$ is a* (finite) linear combination *of vectors in Y iff there is a finite set of vectors $\{y_i\} \subseteq Y$ and scalars $\{\alpha_i\}$ such that*

$$x = \sum_{i=1}^{n} \alpha_i \, y_i$$

Note: The sum is not required to be unique. (Unlike $\oplus$.)

Example 3 *Let $Y = \{(1,0),(1,1),(0,1)\} \subset \mathbb{R}^2$. Then the vector $x = (2,3)$ can be written as $x = 2(1,0) + 3(0,1)$ or as $x = 2(1,1) + 1(0,1)$ or as $x = -1(1,0) + 3(1,1)$.*

Generated Subspace & Span

Theorem 9 *Let* $\emptyset \neq Y \subseteq X$. *Define*

$$V(Y) \stackrel{\Delta}{=} \{\textit{all linear combinations from } Y\}.$$

Then $V(Y)$ *is a subspace of* X *and is called the* subspace generated by Y.

Definition 11 (Span) Y spans X *if and only if* $V(Y) = X$.

Example 4 *Let* $Y = \{(1,0),(1,1),(0,1)\}$. *Then* Y *spans* $\mathbb{R}^2$. *(Exercise.)*

Example 5 *Let* $Z = \{(1,1,0),(1,0,1),(0,1,1),(1,1,1)\}$. *Does the set* Z *span* $\mathbb{R}^3$*?*

Dependence and Independence

Definition 12 (Linear Dependence) *Let* $\{x_1, x_2, \ldots, x_m\}$ *be a nonempty subset of* X*. If there exists a set of scalars* $\{\alpha_i\}$*, not all zero, such that* $\alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_m x_m = 0$*, then* $\{x_1, x_2, \ldots, x_m\}$ *is* linearly dependent*.*

Definition 13 (Linear Independence) *If the nonempty subset* $\{x_1, x_2, \ldots, x_m\}$ *of* X *is not linearly dependent, then* $\{x_1, x_2, \ldots, x_m\}$ *is* linearly independent*.*

Example 6 Y *and* Z *from the previous examples are both linearly dependent.*

Example 7 *Let* $W = \{(1,1,0), (1,0,1), (0,1,1)\}$*. Then* W *is linearly independent.*

More Examples

Example 8 *Let* $V = \{(1,1,0,0),(1,0,1,0),(1,1,1,0)\}$*. Is* V *linearly independent? Does* V *span* $\mathbb{R}^4$*?*

Example 9 *Let* $U = \{(0,0,0),(1,0,0),(0,1,0)\}$*. Is* U *linearly independent?*

Example 10 *Let* $\mathcal{P} = \{1, x, x^2, x^3, \ldots\}$*. Then* $V(\mathcal{P}) = \mathbb{P}$*, the set of all real polynomials; i.e.,* $\mathcal{P}$ *spans* $\mathbb{P}$*. Is* $\mathcal{P}$ *linearly independent? Yes! But how do we show this? Consider*

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

and note that the only n*th degree polynomial with* $n+1$ *roots, is* $p(x) \equiv 0$*.*

(Go to TOC)

9. Linear Independence

Theorem 10 (Uniqueness) *Let* $Y = \{x_1, x_2, \ldots, x_m\}$ *be a linearly independent set of vectors. If* $\sum_{i=1}^{m} \alpha_i x_i = \sum_{i=1}^{m} \beta_i x_i$*, then* $\alpha_i = \beta_i$ *for* $i = 1..m$*.*

Pf. Simple calculation.

Theorem 11 *A set* Y *is linearly dependent if and only if some vector* $x \in Y$ *can be written as a linear combination of other vectors in* Y*.*

- Add any number of vectors to a dependent set, it will still be dependent.
- Add one vector to an independent set, it may or may not stay independent.

Infinite Example

Example 11 *Let* $\mathcal{P} = \{1, x, x^2, x^3, \ldots\}$. *Then* $V(\mathcal{P}) = \mathbb{P}$, *the set of all real polynomials; i.e.,* $\mathcal{P}$ *spans* $\mathbb{P}$. *Is* $\mathcal{P}$ *linearly independent? Yes! But how do we show this? Let* $p(x) \in \mathbb{P}$. *Then, for some* n,

$$p(x) = \sum_{i=0}^{n} \alpha_i x_i = 0.$$

Note that the only n*th degree polynomial with* $n+1$ *roots, is* $p(x) \equiv 0$. *Hence all* α_i *are* 0.

Unique Expression

Theorem 12 (Uniqueness of Expression) *A finite nonempty set Y is linearly independent if and only if, for each nonzero $y \in V(Y)$, there exists a unique subset $\{x_1, \ldots, x_m\}$ of Y and a unique set of scalars $\{\alpha_1, \ldots, \alpha_m\}$ such that $y = \sum_{i=1}^{m} \alpha_i x_i$.*

Assignment:

1. Prove Theorem 11
2. Prove Theorem 12

Theorem 13 *Y is linearly independent if and only if $Z \subsetneq Y$ implies $V(Z) \neq V(Y)$.*

Pf. Exercise.

Basis of a Vector Space

Definition 14 (Hamel basis) *A (finite) set $Y \subseteq X$ is a* Hamel basis *(or just a* basis*) if and only if*

1. *Y is linearly independent*
2. *$V(Y) = X$*

Id est, *Y is a (finite) linearly independent spanning set.*

Theorem 14 *If Y is linearly independent, then Y is a basis for $V(Y)$.*

Pf. Exercise.

Note: The theorem *Every vector space has a basis* is a result of the *Axiom of Choice*.

(Go to TOC)

10. Basis of a Vector Space

Recall:

Definition 15 (Hamel Basis) *A set $Y \subseteq X$ is a* Hamel Basis *(or just a* basis*) if and only if*

1. *Y is linearly independent*
2. *$V(Y) = X$*

Note: The theorem «*Every vector space has a basis*» is a result of the *Axiom of Choice*.

Exempli gratia

- $\{(0,1),(1,2)\}$ is a basis of $\mathbb{R}^2$
- $\{(1,1,0),(1,0,1),(0,1,1)\}$ is a basis of $\mathbb{R}^3$
- $\{(1,1,0),(1,2,0),(2,1,0)\}$ is *not* a basis of $\mathbb{R}^3$

Basis Properties

Theorem 15 (Uniqueness of Scalars) *Let* $\{x_1, x_2, \ldots, x_n\}$ *be a basis for* X. *Then for each vector* $x \in X$, *there is a unique set of scalars* $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ *such that*

$$x = \alpha_1 x_1 + \alpha_2 x_2 + \cdots + \alpha_n x_n$$

Pf. Standard calculation.

Theorem 16 (Maximum Independent Set Size) *Suppose that* $B = \{x_1, x_2, \ldots, x_n\}$ *is a basis of* X *with* n *finite and* $Y = \{y_1, y_2, \ldots, y_m\}$ *is a set of linearly independent vectors. Then* $m \leq n$.

Note: n is *finite* is necessary.

Proof of Theorem 16 - Outline

Proof Outline.

1. Assume $m > n$.
2. Write y_1 as a linear combination of the x_i. At least one coefficient can't be 0, say the coefficient of x_n (reindex x's if necessary).
3. Replace x_n in B with y_1. Show B still is a basis for X.
4. Start over with y_2 and the "new" B. Replace x_{n-1} by y_2.
5. Continue the process until y_n replaces x_1.
6. B - still a basis - now is $\{y_1, y_2, \ldots, y_n\}$.
7. Thus y_{n+1} can be written as as linear combination from B contradicting the linear independence of Y. Hence $m \leq n$.

Dimension

Theorem 17 *If* $B = \{x_1, x_2, \dots, x_N\}$ *is a basis of* X *for some* $N < \infty$*, then every basis of* X *contains exactly* N *vectors.*

Pf. • Let B_1 be a basis with n vectors and B_2 be a basis with m vectors.
• Apply Theorem 16 with B_1 as the basis and B_2 as the linearly independent set. Therefore $m \leq n$.
• Now apply Theorem 16 with B_2 as the basis and B_1 as the linearly independent set. Therefore $n \leq m$.
• Since $m \leq n$ and $n \leq m$, it follows that $m = n$.

(Go to TOC)

11. Dimension of a Vector Space

Definition 16 (Dimension) *If X has a finite basis of n vectors, then X is* finite dimensional *and has* dimension $dim(X) = n$. *If X is not finite dimensional, then X has* infinite dimension *and $dim(X) = \infty$.*

Example 12 *Several standard spaces:*

- $dim(\mathbb{R}^n) = n$
- $dim(\mathbb{P}^n) = n + 1$
- $dim(\mathbb{R}^\infty) = \infty$ *The space of real sequences is large (but it's a "small ∞")*
- $dim(\mathbb{P}) = \infty$ *(another "small ∞," isomorphic to $\mathbb{R}^\infty$)*

Examples

Example 13 *Infinite dimensional spaces*

- $dim(\mathcal{C}[0,1]) = \infty$ *The space of continuous functions on* $[0,1]$ *is very large (a "big* ∞*")*
- $dim(\mathcal{B}(\mathbb{R})) = \infty$ *with* $\mathcal{B}(\mathbb{R}) = \{$*bounded real functions* $\}$
- *Is the following true:*
  Let Z *be an arbitrary set and* X *an arbitrary vector space over* F*. The space of all functions from* Z *to* X*, written* X^Z*, is a vector space over* F *with dimension*
  $dim(X^Z) = dim(X)^{|Z|}$

Basis & Dimension Facts

Basis Facts

- Every vector space has a basis (requires the Axiom of Choice)
- Every linearly independent set can be extended to a basis
- A linearly independent set can be no larger than a basis
- A set containing more vectors than a basis must be linearly dependent
- Any two bases for a vector space contain the same number of vectors (finite dimensional case)
- If X has a set with n linearly independent vectors and every set of $n+1$ vectors is dependent, then $dim(X) = n$
- If Y is a subspace of X, then $dim(Y) \leq dim(X)$.

“Two Out of Three Ain’t Bad”

Theorem 18 *Suppose* X *is a vector space with* $dim(X) = n$ *and* $Y \subseteq X$*. If any two of the following hold, then the third also holds.*

1. Y *spans* X
2. Y *is linearly independent*
3. Y *contains exactly* n *vectors*

Theorem 19 *Suppose that* $dim(X) < \infty$ *and that* $X = Y \oplus Z$*. Then* $dim(X) = dim(Y) + dim(Z)$*.*

Nota Bene: Recall that $\oplus$ is the “interior analogue” of $\times$ and that if $X = Y \times Z$, then $dim(X) = dim(Y) \times dim(Z)$.

“Sum of Dimensions” Proof

Proof of Theorem 19 (3.3.43).

Since $dim(X) < \infty$, so are $dim(Y)$ and $dim(Z)$. Therefore there are bases of Y and Z: $\mathcal{B}_Y = \{y_1, \ldots, y_n\}$ and $\mathcal{B}_Z = \{z_1, \ldots, z_m\}$. Set $\mathcal{B} = \mathcal{B}_Y \cup \mathcal{B}_Z$. Let

$$0 = \sum_{i=1}^{n} \alpha_i y_i + \sum_{i=1}^{m} \beta_i z_i$$

be a linear combination from $\mathcal{B}$. Since representation of vectors is unique in $X = Y \oplus Z$, we have that $0 = \sum_{i=1}^{n} \alpha_i y_i$ and $0 = \sum_{i=1}^{m} \beta_i z_i$ Therefore $0 = \alpha_i = \beta_j$ for all i and j as $\mathcal{B}_Y$ and $\mathcal{B}_Z$ are independent. I.e., $\mathcal{B}$ is linearly independent.

Since $X = Y \oplus Z$, it is clear that $\mathcal{B}$ spans X. Hence, $|\mathcal{B}| = n + m = dim(X)$.

(Go to TOC)

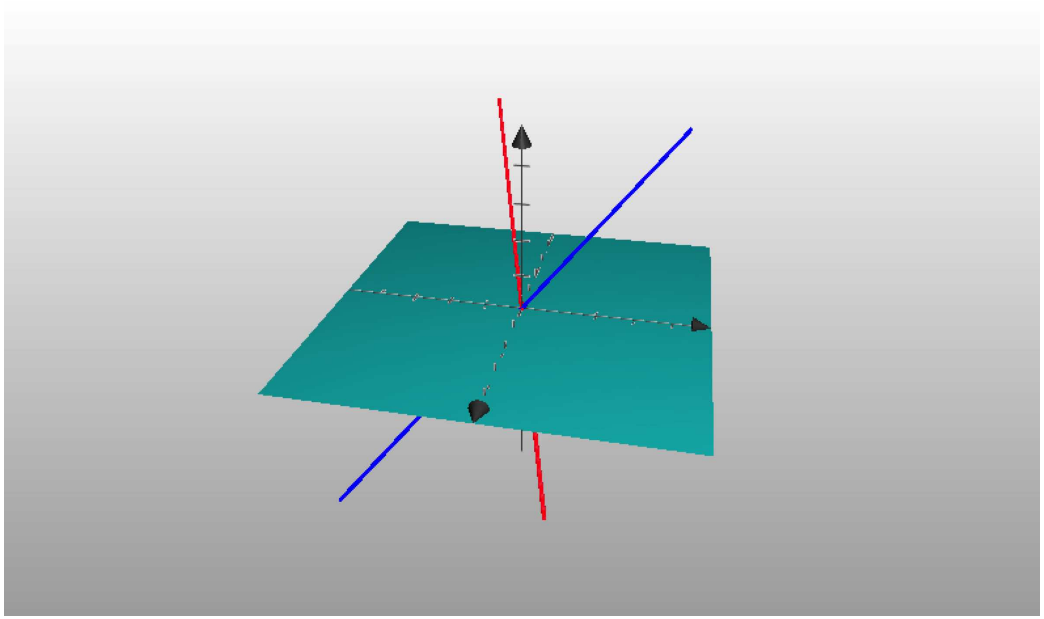
12. Subspaces and Direct Sums

The theorem *Every linearly independent set can be extended to a basis* has:

Corollary 20 *Suppose X is an n-dimensional vector space with an m-dimensional subspace Y. Then there exists a subspace Z of dimension $(n-m)$ such that $X = Y + Z$.*

Pf. (Sketch) Take bases $\mathcal{B}_X$ for X and $\mathcal{B}_Y$ for Y. Eliminate the portion of $\mathcal{B}_X$ dependent on $\mathcal{B}_Y$. The remaining vectors form a basis for Z.

Note: Z need not be unique. (Z= red or blue)



Linear Transformations

Definition 17 (Linear Transformation) *A mapping T from a vector space X into a vector space Y, both spaces over the field F, is a* linear transformation, *written as $T \in L(X, Y)$, if and only if for all $x \in X$, $y \in Y$, and $\alpha \in F$, we have*

1. $T(x + y) = T(x) + T(y)$
2. $T(\alpha x) = \alpha T(x)$

A nonlinear transformation *is a mapping that is not linear.*

Theorem 21 (Superposition Principle) $T \in L(X, Y)$ *if and only if*

$$T\left(\sum_{i=1}^{m} \alpha_i x_i\right) = \sum_{i=1}^{m} \alpha_i T(x_i)$$

Examples of Linear Transformations

Example 14

- $T : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $T([x, y]) = [2x + 3y, x - y]$
- $T : \mathbb{R}^2 \to \mathbb{R}^1$ *by* $T([x, y]) = [x]$
- $D : \mathbb{P}^n \to \mathbb{P}^{n-1}$ *by* $D(p) = \frac{d}{dx} p(x)$
- $I : \mathcal{C}[0, 1] \to \mathbb{R}$ *by* $I(f) = \int_0^1 f(t)\, dt$
- *Let* $k \in \mathcal{C}[a, b] \times \mathcal{C}[a, b]$ *such that for any* $x \in \mathcal{C}[a, b]$,

$$\widehat{x}(s) = \int_a^b x(t)k(s, t)\, dt \in \mathcal{C}[a, b]$$

Then $\widehat{\cdot} : \mathcal{C}[a, b] \to \mathcal{C}[a, b]$ *is a linear transformation.*[a]

[a] Fredholm Integral Equation of the First Type *or a* kernel transform

(Go to TOC)

13. Examples of Linear Transformations

Example 15

- *Let* $\mathcal{L}_1^c = \{f : \mathbb{R} \to \mathbb{C} \mid f \in \mathcal{C}(\mathbb{R})$ *and* $\int_{\mathbb{R}} |f| < \infty\}$*. Now define the* Fourier transform $\mathcal{F}(f) \in \mathcal{L}_1^c$ *by*

$$\mathcal{F}(f)(s) = \int_{\mathbb{R}} f(t)\, e^{-ist}\, dt$$

  Then $\mathcal{F} : \mathcal{L}_1^c \to \mathcal{L}_1^c$ *is a linear transformation.*

- *Let* $z \in \mathbb{C}$*. Then* $\bar{z} =$ *(the complex conjugate of* z*) is a* nonlinear transformation*.*

- *Let* $|\cdot|$ *be the absolute value function on* $\mathbb{R}$*. Is* $|\cdot|$ *a linear transformation from* $\mathbb{R}$ *to* $\mathbb{R}$*?*

Null Space and Range Space

Definition 18 *Let* $T \in L(X, Y)$. *Then the*

1. null space $\mathcal{N}(T)$ *(or* kernel $\ker(T)$*) is the set*

$$\mathcal{N}(T) = \{x \in X \mid T(x) = 0\},$$

2. range space $\mathcal{R}(T)$ *(or* image space*) is the set*

$$\mathcal{R}(T) = \{y \in Y \mid y = T(x) \text{ for some } x \in X\} = T(X).$$

Theorem 22 *Let* $T \in L(X, Y)$. *Then*

1. $\mathcal{N}(T)$ *is a subspace of* X,
2. $\mathcal{R}(T)$ *is a subspace of* Y.

Pf. Exercise (3.4.20)

Range & Dimension

Theorem 23 *If* $T \in L(X, Y)$, *then* $\dim(\mathcal{R}(T)) \le \dim(X)$*.*

Pf. Assume $X \neq \{0\} \neq \mathcal{R}(T)$, otherwise the result is trivial. Set $n = dim(X) > 0$. Choose $\{y_1, \ldots, y_{n+1}\} \subseteq \dim(\mathcal{R}(T))$. For each i, find x_i such that $T(x_i) = y_i$. Since $\dim(X) = n$, we know that there are scalars α_i so that

$$\alpha_1 x_1 + \cdots + \alpha_{n+1} x_{n+1} = 0$$

Applying T to this linear combination yields

$$\alpha_1 y_1 + \cdots + \alpha_{n+1} y_{n+1} = 0$$

Since the y_i were arbitrary, every subset of $\dim(\mathcal{R}(T))$ with $n+1$ vectors is linearly dep. Thence $\dim(\mathcal{R}(T)) \le n$.

(Go to TOC)

14. The Dimension Theorem

Theorem 24 *The inverse image of a basis under a linear transformation is linearly independent. I.e., Let* $T \in L(X, Y)$ *and let* $\mathcal{B}_Y = \{y_i\}$*. For each* i*, choose an* x_i *such that* $T(x_i) = y_i$*. Then the set* $\{x_i\}$ *is linearly independent.*

Pf. Exercise (3.4.24)

Theorem 25 (The Dimension Theorem) *Let* $T \in L(X, Y)$ *with* $\dim(X) < \infty$*. Then*

$$\dim(\mathcal{R}(T)) + \dim(\mathcal{N}(T)) = \dim(X).$$

Pf. Set $\dim(X) = n$ and $\dim(\mathcal{N}(T)) = s$ and set $r = n - s$. (*Need to show*: $\dim(\mathcal{R}(T)) = r = n - s$.)

The Dimension Theorem Proof

Pf. Find a basis for $\mathcal{N}(T)$ labeling the vectors $\{e_1, \ldots, e_s\}$. Extend this set to a basis for X by adding r vectors to have $\mathcal{B} = \{x_1, \ldots, x_r, e_1 \ldots, e_s\}$. Since $\mathcal{B}$ is a basis, then $T(\mathcal{B})$ spans $\mathcal{R}(T)$. Since $T(e_i) = 0$, then $T(\{x_1, \ldots, x_r\})$ spans $\mathcal{R}(T)$. Set $y_i = T(x_i)$; so $\{y_1, \ldots, y_r\}$ spans $\mathcal{R}(T)$. Suppose a linear combination $\alpha_1 y_1 + \cdots + \alpha_r y_r = 0$. Then because $\sum_r \alpha_i T(x_i) = T(\sum_r \alpha_i x_i)$, we have that $\sum_r \alpha_i x_i \in \mathcal{N}(T)$, thus $\sum_r \alpha_i x_i = \sum_s \gamma_i e_i$ which can be written as

$$\alpha_1 x_1 + \cdots + \alpha_r x_r - \gamma_1 e_1 - \cdots - \gamma_s e_s = 0$$

which implies each $\alpha_i = 0$. Hence $\dim(\mathcal{R}(T)) = r$.

– *What about the cases* $s = 0$ *and* n? (Group-Project time!)

(Go to TOC)

“Your Turn”

The Setup. Define $\mathcal{D} : \mathbb{R}^4 \to \mathbb{R}^4$ by

$$\mathcal{D}\left([\,x_1, x_2, x_3, x_4\,]\right) = [\,x_2, 2x_3, 3x_4, 0\,]$$

The Project.

1. Is $\mathcal{D}$ a linear transformation?
2. What is $\mathcal{R}(T)$?
3. Find $\dim(\mathcal{R}(T))$.
4. What is $\mathcal{N}(T)$?
5. Find $\dim(\mathcal{N}(T))$.
6. Calculate $\dim(\mathcal{R}(T)) + \dim(\mathcal{N}(T))$.

(Go to TOC)

15. Rank & Nullity

Definition 19 (Rank and Nullity of a Linear Transformation) *Let* $T \in L(X, Y)$*.*

- *The* rank ρ *of* T *is the dimension of the range space;* $\rho(T) = \dim(\mathfrak{R}(T))$
- *The* nullity ν *of* T *is the dimension of the nullspace;* $\nu(T) = \dim(\mathfrak{N}(T))$

Corollary 26 (Fundamental Theorem of Linear Algebra) *Let* $T \in L(X, Y)$ *where* $\dim(X) = n$*. Then*

$$\rho(T) + \nu(T) = n$$

Pf. $\checkmark$

"Affine Nullspace"

Corollary 27 *Let* $T \in (X, Y)$ *where* $\dim(X) < \infty$, *and let* $\mathcal{B} = \{x_1, \ldots, x_s\}$ *be a basis for* $\mathfrak{N}(T)$ *so that* $\dim(\mathfrak{N}(T)) = s$. *Then*

1. *a vector* $x \in X$ *satisfies* $T(x) = 0$ *iff there is a unique set of scalars* α_i s.t. $x = \sum_{i=1}^{s} \alpha_i x_i$,
2. *a vector* $y_0 \in Y$ *is in* $\mathfrak{R}(T)$ *iff there is at least one vector* $x \in X$ s.t. $y_0 = T(x)$,
3. *if vectors* $x_0 \in X$ *and* $y_0 \in Y$ *are s.t.* $T(x_0) = y_0$, *then* $x \in X$ *satisfies* $T(x) = y_0$ *iff there is a unique set of scalars* β_i s.t. $x = x_0 + \sum_{i=1}^{s} \beta_i x_i$.

Pf. ✓

Inverses

Theorem 28 *Let* $T \in L(X,Y)$.

1. T^{-1} *exists iff* $T(x) = 0$ *implies* $x = 0$*; i.e.,* $\mathfrak{N}(T) = \{0\}$.
2. *If* T^{-1} *exists, then* $T^{-1} \in L(\mathfrak{R}(T), X)$.

Pf. $1.$ ($\Leftarrow$) Assume $\mathfrak{N}(T) = \{0\}$. Then $T(x_1) = T(x_2) \Leftrightarrow T(x_1) - T(x_2) = 0 \Leftrightarrow T(x_1 - x_2) = 0 \Leftrightarrow x_1 - x_2 \in \mathfrak{N}(T) \Leftrightarrow x_1 = x_2$.
($\Rightarrow$) Now assume that T^{-1} exists and that $T(x) = 0$. Since $T(0) = 0$, then $T(x) = T(0)$. Whence $x = 0$.

$2.$ Assume that T is nonsingular and that $T(x_1) = y_1$, $T(x_2) = y_2$. Then $T^{-1}(y_1 + y_2) = T^{-1}(T(x_1) + T(x_2)) = T^{-1}(T(x_1 + x_2)) = x_1 + x_2 = T^{-1}(y_1) + T^{-1}(y_2)$. For $\alpha \in F$, $T^{-1}(\alpha y_1) = T^{-1}(\alpha T(x_1)) = T^{-1}(T(\alpha x_1) = \alpha x_1 = \alpha T^{-1}(y_1)$.

Examples

Example Set 16

- *Let* $T([a,b]) = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} [a,b]$. *Show* T *is nonsingular.*
- *Let* $S([a,b]) = \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} [a,b]$. *Show* T *is singular.*
- $\mathcal{D} : \mathbb{P} \to \mathbb{P}$ *defined by* $\mathcal{D}(p) = \dfrac{dp}{dx}$ *is singular.*
- *Is* $\mathcal{I} : \mathbb{P} \to \mathbb{P}$ *defined by* $\mathcal{I}(p) = \int p\,dx$ *nonsingular?*
- *Is* $T([a,b]) = [a+b, 0, a-b, 0, 0]$ *invertible?*

(Go to TOC)

Project Solution

The Group-Project solution is much easier when looking at the spaces from a different “dimension.”

The Setup. Define $\mathcal{D} : \mathbb{R}^4 \to \mathbb{R}^4$ by

$$\mathcal{D}\left([\,x_1, x_2, x_3, x_4\,]\right) = [\,x_2, 2x_3, 3x_4, 0\,]$$

A Solution. Consider $\mathcal{T} : \mathbb{P}^3 \to \mathbb{P}^3$ with $\mathcal{T}(p) = p'$. $(\mathbb{P}^3 \cong \mathbb{R}^4)$

- $\mathfrak{R}(T) = \{$polynomials of degree $2\} \cong \mathbb{R}^3$
- $\mathfrak{N}(T) = \{$constant polynomials$\} \cong \mathbb{R}^1$
- $4 = 3 + 1 \Rightarrow \mathbb{R}^4 \cong \mathbb{P}^3 = \left\{p \in \mathbb{P}^3 \mid p(0) = 0\right\} \oplus \mathfrak{N}(T)$

(Go to TOC)

16. Singular and Nonsingular Examples

Example Set 17

- *Let* $T([a,b]) = \begin{bmatrix} 1 & 2 \\ 1 & 0 \end{bmatrix} [a,b]$. *Show* T *is nonsingular.*
- *Let* $S([a,b]) = \begin{bmatrix} 6 & 3 \\ 2 & 1 \end{bmatrix} [a,b]$. *Show* T *is singular.*
- $\mathcal{D} : \mathbb{P} \to \mathbb{P}$ *defined by* $\mathcal{D}(p) = \dfrac{dp}{dx}$ *is singular.*
- *Is* $\mathcal{I} : \mathbb{P} \to \mathbb{P}$ *defined by* $\mathcal{I}(p) = \int p\, dx$ *nonsingular?*
- *Is* $T([a,b]) = [a+b, 0, a-b, 0, 0]$ *invertible?*

"Inverse Results"

Theorem 29 *Let* $T \in L(X,Y)$ *with* $\dim(X) < \infty$. *Then* T *is invertible if and only if* $\rho(T) = \dim(X)$. T *is said to have "full rank."*

Pf. ✓

Theorem 30 *Let* $T \in L(X,Y)$ *with* $\dim(X) = \dim(Y) = n$ *where* $n < \infty$. *Then* T *is invertible if and only if* $\mathfrak{R}(T) = Y$.

Pf. $(\Rightarrow)$ T invertible implies that $\dim(\mathfrak{R}(T)) = n = \dim(Y)$. Since $\mathfrak{R}(T)$ is a subspace of Y, then $\mathfrak{R}(T) = Y$.

$(\Leftarrow)$ Choose a basis $\mathcal{B} = \{y_1, \ldots, y_n\}$ for $\mathfrak{R}(T) = Y$. Then, since $T^{-1}(\mathcal{B})$ is an independent set of size n, it forms a basis for X. Hence the only set of scalars for which $\sum_i \alpha_i x_i = 0$ is $\alpha_i = 0$. Whence $\mathfrak{N}(T) = \{0\}$, so T is invertible.

Collected Results, I

Theorem 31 (Invertible Linear Transformations) *Let X and Y be vector spaces over F and let $T \in L(X, Y)$. TFAE:*

1. T *is invertible or nonsingular*
2. T *is injective or* 1–1
3. $T(x) = 0$ *implies* $x = 0$; *i.e.,* $\mathfrak{N}(T) = \{0\}$
4. *For each* $y \in Y$, $\exists$ *a unique* $x \in X$ *such that* $T(x) = y$
5. *If* $T(x_1) = T(x_2)$, *then* $x_1 = x_2$
6. *If* $x_1 \neq x_2$, *then* $T(x_1) \neq T(x_2)$,

If X is finite dimensional, then TFAE:

7. T *is injective*
8. $\rho(T) = \dim(X)$

Collected Results, II

Theorem 32 (Surjective Linear Transformations) *Let X and Y be vector spaces over F and let $T \in L(X, Y)$. TFAE:*

1. *T is surjective or onto*
2. *For $y \in Y$, there is at least one $x \in X$ such that $T(x) = y$*

If X and Y are finite dimensional, then TFAE:

3. *T is surjective*
4. *$\rho(T) = \dim(Y)$*

Pf. ✓

Collected Results, III

Theorem 33 (Bijective Linear Transformations) *Let X and Y be vector spaces over F and let $T \in L(X,Y)$. TFAE:*

1. *T is bijective or onto*
2. *For $y \in Y$, there is a unique $x \in X$ such that $T(x) = y$*

If X and Y are finite dimensional, then TFAE:

3. *T is surjective*
4. *$\rho(T) = \dim(X) = \dim(Y)$*

Theorem 34 (Common Finite Dimension) *Let X and Y be vector spaces over F with finite dimension n and $T \in L(X,Y)$. Then*

$$T\text{:injective} \Leftrightarrow T\text{:surjective} \Leftrightarrow T\text{:bijective} \Leftrightarrow T\text{:invertible}$$

Transformation Spaces

Definition 20 *For S and T in $L(X,Y)$ and α in F, define*

1. $S+T$ *by* $(S+T)(x) \triangleq S(x)+T(x)$
2. $\alpha\, S$ *by* $(\alpha\, S)(x) \triangleq \alpha\, S(x)$
3. $S \circ T$ *by* $(S \circ T)(x) \triangleq S\,(T(x))$

Theorem 35 $L(X,Y)$ *is a vector space over* F *(using 1 & 2)*

Theorem 36 $L(X,X)$ *is an associative algebra with identity over* F *(using 1, 2, & 3, and identity* $I(x)=x$*)*

17. Transformation Spaces

Definition 21 *For S and T in $L(X,Y)$ and α in F, define*

1. $S+T$ *by* $(S+T)(x) \stackrel{\Delta}{=} S(x)+T(x)$
2. $\alpha\, S$ *by* $(\alpha\, S)(x) \stackrel{\Delta}{=} \alpha\, S(x)$
3. ST *by* $(ST)(x) \stackrel{\Delta}{=} S\,(T(x))$ *when* $\mathrm{range}(T) \subseteq \mathrm{dom}(S)$

Theorem 37 *Let $S, T, U \in L(X,X)$. Then*

1. *If $ST = US = I$, then S is bijective and $S^{-1} = T = U$.*
2. *If S is bijective, then $(S^{-1})^{-1} = S$.*
3. *If S and T are bijective, then $(ST)^{-1} = T^{-1}S^{-1}$.*
4. *If S is bijective and $\alpha \neq 0$, then $(\alpha S)^{-1} = (1/\alpha)\cdot S^{-1}$.*

Polynomials of Transforms

Theorem 38 *$L(X,X)$ is an associative algebra[a] with identity over F (using 1, 2, & 3, and identity $I(x)=x$). $L(X,X)$ is usually noncommutative.*

Definition 22 (Powers of Transforms) *Let $T \in L(X,X)$. Then set $T^0 = I$ and, for $n > 0$, define $T^{(n)} \triangleq T \cdot T^{(n-1)}$ and $T^{(-n)} \triangleq \left(T^{-1}\right)^n$.*

Definition 23 *Let $p \in \mathbb{P}^n$, so that $p(\lambda) = a_0 + a_1\lambda + \cdots + a_n\lambda^n$. For $T \in L(X,X)$, define*

$$p(T) = a_0\, I + a_1 T + \cdots + a_n T^n = \sum_{i=0}^{n} \alpha_i T^i.$$

[a] "Vector space plus multiplication." See pg. 56 and 104 of the text.

Finite Dimension Structure Theorem

Definition 24 X *is* isomorphic *to* Y, *written* $X \cong Y$, *if and only if there is a bijection* $T \in L(X,Y)$.

Theorem 39 (Structure Theorem) *Every* n*-dimensional vector space* X *over the field* F *is isomorphic to* F^n.

Pf. Choose a basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ for X. Then define $T \in L(X, F^n)$ by

$$T\left(\sum_{i=1}^{n} \alpha_i e_i\right) = [\alpha_1, \alpha_2, \ldots, \alpha_n]$$

Corollary 40 *Let* F *be a field and* n *be a positive integer. There is exactly one vector space of dimension* n *over* F.

(Go to TOC)

18. Linear Functionals

Definition 25 *Let X be a vector space over F. Then $f \in L(X, F)$ is called a* linear functional*.*

Example Set 18

- *Let $f \in \mathcal{C}[a, b]$. Then $F(f) = \int_a^b f(t)\, dt$ is a linear functional.*
- *Let $f \in \mathcal{C}[a, b]$ and choose $k \in \mathcal{C}[a, b]$. Then $F_k(f) = \int_a^b f(t)k(t)\, dt$ is a linear functional.*
- *Let $f \in \mathcal{C}[a, b]$ and $x_0 \in [a, b]$. Is $\frac{df}{dt}(x_0)$ a linear functional?*
- *Let F be a field. The mappings $\mathrm{proj}_i : F^n \to F$ for $i = 1..n$ given by $\mathrm{proj}_i \left([\alpha_1, \alpha_2, \ldots, \alpha_n]\right) = \alpha_i$ are linear functionals. $\phi = \sum \alpha_i \, \mathrm{proj}_i$ is also a linear functional.*

Vector Space of Linear Functionals

Definition 26 *Let* X *be a vector space over* F*. Define* $X^f = L(X, F)$*. When* $f \in X^f$ *is evaluated at the vector* $x \in X$*, we use the notation* $f(x) \stackrel{\Delta}{=} \langle x, f \rangle$*. Using* x' *in place of* $f \in X^f$*, we see*

$$\begin{aligned}(f_1 + f_2)(x) &= \langle x, x_1' + x_2' \rangle \stackrel{\Delta}{=} \langle x, x_1' \rangle + \langle x, x_2' \rangle \\ &= f_1(x) + f_2(x)\end{aligned}$$

and

$$\begin{aligned}(\alpha f)(x) &= \langle x, \alpha x' \rangle \stackrel{\Delta}{=} \alpha \langle x, x' \rangle \\ &= \alpha f(x)\end{aligned}$$

Theorem 41 $X^f = L(X, F)$ *is a vector space over* F *called the* algebraic conjugate *of* X.

Algebraic Conjugate Basis

Theorem 42 *Let X be a vector space with basis $\mathcal{B} = \{e_1, \ldots, e_n\}$ and let $\{\alpha_1, \ldots, \alpha_n\}$ be a set of arbitrarily chosen scalars. Then there is a unique linear functional $x' \in X^f$ such that $\langle e_i, x' \rangle = \alpha_i$ for $i = 1..n$.*

Pf. $(\exists)$ For every $x \in X$, we have unique scalars ξ_i such that $x = \sum_n \xi_i e_i$. Define $x' \in X^f$ by $\langle x, x' \rangle = \sum_n \alpha_i \xi_i$. If $x = e_i$ for some i, then $\xi_i = 1$ and $\xi_j = 0$ for every $j \neq i$. Hence $\langle x, x' \rangle = \alpha_i$; i.e., $\langle e_i, x' \rangle = \alpha_i$.
$(!)$ Suppose $\langle e_i, x'_1 \rangle = \alpha_i$ and $\langle e_i, x'_2 \rangle = \alpha_i$ for $i = 1..n$. Then $\langle e_i, x'_1 \rangle - \langle e_i, x'_2 \rangle = 0$ for $i = 1..n$, and so $\langle e_i, x'_1 - x'_2 \rangle = 0$ for $i = 1..n$. This implies that $x'_1 = x'_2$.

Definition 27 (Kronecker Delta) *Set* $\delta_{ij} = \begin{cases} 1 & \textit{if } i = j \\ 0 & \textit{if } i \neq j \end{cases}$.

(Go to TOC)

19. Conjugate Dimension Theorem

Theorem 43 *Let* X *be a finite dimensional vector space with basis* $\mathcal{B} = \{e_1, \ldots, e_n\}$. *Then there exists a unique basis* $\mathcal{B}' = \{e'_1, \ldots, e'_n\}$ *for* X^f *such that* $\langle e_i, e'_j \rangle = \delta_{ij}$; *we call* $\mathcal{B}'$ *the* dual basis *of* $\mathcal{B}$. *Further* $\dim(X) = n = \dim(X^f)$.

Pf. There exists a unique set of linear functionals $\mathcal{B}' = \{e'_j\}$ such that $\langle e_i, e'_j \rangle = \delta_{ij}$ for $i, j = 1..n$ which are found by applying the previous theorem to the sets
$A_j = \{\delta_{ij} | j = 1..n\}$.

$(\mathcal{B}'$ is linearly independent$)$ Since $\sum \beta_i e'_i = 0$ implies

$$0 = \left\langle e_j, \sum_i \beta_i e'_i \right\rangle = \sum_i \beta_i \langle e_j, e'_i \rangle = \sum_i \beta_i \delta_{ij} = \beta_j$$

Conjugate Dimension Theorem, II

(Pf.) $(\mathcal{B}'$ spans $X^f)$ Let $x' \in X^f$ and define $\alpha_i = \langle e_i, x' \rangle$. (This form is often called a *projection*.) For $x \in X$, there are scalars so that $x = \sum_i \xi_i e_i$. Then

$$\langle x, x' \rangle = \left\langle \sum_i \xi_i e_i, x' \right\rangle = \sum_i \langle \xi_i e_i, x' \rangle = \sum_i \xi_i \langle e_i, x' \rangle = \sum_i \xi_i \alpha_i$$

It also follows that $\langle x, e'_j \rangle = \sum_i \xi_i \langle e_i, e'_j \rangle = \xi_j$. Combine these two results to obtain

$$\langle x, x' \rangle = \sum_i \alpha_i \langle x, e'_i \rangle = \left\langle x, \sum_i \alpha_i e'_i \right\rangle$$

which gives us $x' = \sum_i \alpha_i e'_i$.

(Go to TOC)

20. Algebraic Transpose

Definition 28 (Algebraic Transpose) *Let* $S \in L(X, Y)$. *Then* $S^T : Y^f \to X^F$ *given by* $\langle x, S^T y' \rangle = \langle Sx, y' \rangle$ *is the* algebraic transpose *of* S.

Example 19 *Let* $X = \mathbb{R}^3$ *and* $Y = \mathbb{R}^2$. *Define* $S \in L(\mathbb{R}^3, \mathbb{R}^2)$ *by* $y = S(x) = \left[\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ *and* $y' \in Y^f$ *by* $\langle y, y' \rangle = [1, 1] \left[\begin{smallmatrix} y_1 \\ y_2 \end{smallmatrix}\right]$. *Then* $x' = S^T(y')$ *is found by*

$\langle x, x' \rangle = \langle x, S^T y' \rangle = \langle Sx, y' \rangle$	$x'(x) = (S^T(y'))(x)$
$\langle x, x' \rangle = \langle \left[\begin{smallmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \end{smallmatrix}\right] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, y' \rangle$	$x'(x) = y'(S(x))$
$\langle x, x' \rangle = \langle \left[\begin{smallmatrix} x_1 + x_2 \\ x_2 + x_3 \end{smallmatrix}\right], y' \rangle$	$x'(x) = y'(\left[\begin{smallmatrix} x_1 + x_2 \\ x_2 + x_3 \end{smallmatrix}\right])$
$\langle x, x' \rangle = x_1 + 2x_2 + x_3$	$x'(x) = x_1 + 2x_2 + x_3$

The Space of Algebraic Transposes

Theorem 44 *Let* S^T *be the algebraic transpose of* S *where* $S \in L(X,Y)$. *Then* $S^T \in L(Y^f, X^f)$.

Pf. (Calculation.)

1. $S^T(y_1' + y_2') = S^T(y_1') + S^T(y_2')$:

$$\langle x, S^T(y_1' + y_2')\rangle = \langle Sx, (y_1' + y_2')\rangle = \langle Sx, y_1'\rangle + \langle Sx, y_2'\rangle$$
$$= \langle x, S^T y_1'\rangle + \langle x, S^T y_2'\rangle$$

2. $S^T(\alpha y') = \alpha\, S^T(y')$:

$$\langle x, S^T(\alpha y')\rangle = \langle Sx, \alpha y'\rangle = \alpha\langle Sx, y'\rangle$$
$$= \alpha\langle x, S^T(y')\rangle = \langle x, \alpha\, S^T(y')\rangle$$

Algebra of Algebraic Transposes

Theorem 45 *Let I be the identity transform of $L(X, X)$. Then I^T is the identity transform of $L(X^f, X^f)$.*

Theorem 46 *Let 0 be the zero transform of $L(X, Y)$. Then 0^T is the zero transform of $L(Y^f, X^f)$.*

Theorem 47 *Let $R, S \in L(X, Y)$ and $T \in L(Y, Z)$ and let R^T, S^T, and T^T be the respective transposes. Then*

1. $(R + S)^T = R^T + S^T$
2. $(TS)^T = S^T T^T$

Exercise 3.52.32 *(Pg. 113.) Prove the theorems.*

(Go to TOC)

21. Bilinear Functionals

Recall: We have $\overline{a+bi} = a - bi$ for any complex number.

Definition 29 (Conjugate Functional) *Let X be a vector space over $\mathbb{C}$. A mapping $g : X \to \mathbb{C}$ is a* conjugate functional *iff $g(\alpha_1 x_1 + \alpha_2 x_2) = \overline{\alpha_1}\, g(x_1) + \overline{\alpha_2}\, g(x_2)$ for all $x_i \in X$ and $\alpha_i \in \mathbb{C}$*

Definition 30 (Bilinear Form) *Let X be a vector space over $\mathbb{C}$. A mapping $g : X \times X \to \mathbb{C}$ is a* bilinear form *or* bilinear functional *iff for all x, x_i and $y, y_i \in X$ and $\alpha_i, \beta_i \in \mathbb{C}$*

1. $g(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1\, g(x_1, y) + \alpha_2\, g(x_2, y)$
2. $g(x, \beta_1 y_1 + \beta_2 y_2) = \overline{\beta_1}\, g(x, y_1) + \overline{\beta_2}\, g(x, y_2)$

That is, g is linear in the first variable *and* conjugate linear in the second variable*.*

Examples

Example Set 20

1. *Let* $X = \mathbb{C}^2$ *and* g *be given by*

$$g(z_1, z_2) = \mathfrak{Re}(z_1)\mathfrak{Re}(z_2) + \mathfrak{Im}(z_1)\mathfrak{Im}(z_2).$$

2. *Let* $X = \mathbb{R}^2$ *and* h *be given by*

$$h(x, y) = \vec{x} \cdot \vec{y} = x_1 y_1 + x_2 y_2.$$

3. *Let* X *be a vector space over* $\mathbb{C}$ *and let* $P, Q \in X^f$. *Then* $k(x_1, x_2) = P(x_1)\overline{Q(x_2)}$ *is a bilinear functional.*

4. *The conjugate of a bilinear functional is also a bilinear functional. I.e.,* $h(x, y) = \overline{g(x, y)}$.

Definitions

Definition 31 *Let X be a vector space over $\mathbb{C}$ and g be a bilinear functional on X. Then for all $x, y \in X$,*

- g *is* symmetric *iff* $g(x,y) = \overline{g(y,x)}$.
- g *is* positive *iff* $g(x,x) \geq 0$.
- g *is* strictly positive *iff* $g(x,x) > 0$ *whenever* $x \neq 0$.
- $\tilde{g}(x) = g(x,x)$ *is the* quadratic form induced by g.

Example 21 *For $h : \mathbb{R}^2 \to \mathbb{R}$ of Example Set 20, No 2, the induced quadratic form is $\tilde{h}(x) = \tilde{h}([x_1, x_2]) = x_1^2 + x_2^2$.*

(Go to TOC)

22. Quadratic Forms & Inner Products

Theorem 48 *Let g be a bilinear functional. Then*

$$\frac{g(x,y)+g(y,x)}{2} = \tilde{g}\left(\frac{x+y}{2}\right) - \tilde{g}\left(\frac{x-y}{2}\right)$$

Theorem 49 (Polarization) *Let X be a vector space over $\mathbb{C}$ and g be a bilinear functional on X. Then*

$$g(x,y) = \left[\tilde{g}\left(\frac{x+y}{2}\right) - \tilde{g}\left(\frac{x-y}{2}\right)\right] + i\left[\tilde{g}\left(\frac{x+iy}{2}\right) - \tilde{g}\left(\frac{x-iy}{2}\right)\right]$$

Pfs. ✓

"Symmetry is Real"

Theorem 50 *Let* g *and* h *be bilinear functionals on the complex vector space* X. *If* $\tilde{g} = \tilde{h}$, *then* $g = h$.

Theorem 51 *A bilinear functional* g *on a complex vector space* X *is symmetric iff* $\tilde{g}$ *is real.*

Pf. $(\Rightarrow)$ Let g be symmetric, then $g(x,y) = \overline{g(y,x)}$ so that $\tilde{g}(x) = \overline{\tilde{g}(x)}$. Hence $\tilde{g}$ is real.[a]

$(\Leftarrow)$ If $\tilde{g}$ is real, set $h(x,y) = \overline{g(y,x)}$. Then $\tilde{h}(x) = \overline{\tilde{g}(x,x)} = g(x,x) = \tilde{g}(x)$; i.e., $\tilde{h} = \tilde{g}$. By the previous theorem, $h = g$, and hence $g(x,y) = \overline{g(y,x)}$. That is g is symmetric.

---

[a] $z = \bar{z} \;\Rightarrow\; x + iy = x - iy \;\Rightarrow\; y = 0 \;\Rightarrow\; z \in \mathbb{R}$.

Inner Product

Ex. Work through example 3.6.18 on pg. 117.

Definition 32 (Inner Product) *A bilinear functional g is an* inner product *iff*

1. *g is strictly positive* $g(x, x) > 0$ *whenever* $x \neq 0$
2. *g is symmetric* $g(x, y) = \overline{g(y, x)}$

Definition 33 (Inner Product) *(Alternate Definition) A function* $(\cdot, \cdot) : X \times X \to \mathbb{C}$ *is an* inner product *iff*

1. $(x, x) > 0$ *whenever* $x \neq 0$ *and* $(0, 0) = 0$
2. $(x, y) = \overline{(y, x)}$
3. $(\alpha x + \beta y, z) = \alpha(x, z) + \beta(y, z)$
4. $(x, \alpha y + \beta z) = \bar{\alpha}(x, y) + \bar{\beta}(x, z)$

Inner Product Space

Definition 34 *A complex vector space with an inner product is an* inner product space. *A subspace of an inner product space with the restricted inner product is an* inner product subspace.

Definition 35 *Let X be an inner product space. Two vectors x and y are* orthogonal, *written as $x \perp y$, iff $(x, y) = 0$. If x is orthogonal to every vector in a set $A \subseteq X$, then $x \perp A$.*

Example Set 22

1. *Let $X = \mathbb{R}^2$ and let $(x, y) = x_1 y_1 + x_2 y_2$. Then $\{X; (\cdot, \cdot)\}$ is a real inner product space.*
2. *Let $X = \mathbb{C}^n$ and let $(u, v) = \sum_n u_i \overline{v_i}$. Then $\{X; (\cdot, \cdot)\}$ is a complex inner product space.*

(Go to TOC)

23. Inner Product Space Examples

Example 23 *Let* $X = \mathcal{C}_{\mathbb{C}}[0,1]$ *and set* $(f,g) = \int_0^1 f(t)\overline{g(t)}\,dt.$

1. $(t^2+it, 1-it) = \int_0^1 (t^2+it)(1+it)dt = \frac{3}{4}\,i$
2. $(t^2+it, 36t+(2t-25)i) = 0$, *thence it follows that* $(t^2+it) \perp (36t+(2t-25)i)$.
3. $(e^{2\pi kit}, e^{2\pi nit}) = \int_0^1 e^{2\pi(k-n)it}dt = \frac{i}{2\pi(n-k)}\, e^{-2\pi i(n-k)t}\Big|_0^1$
   So $(e^{2\pi kit}, e^{2\pi nit}) = \delta_{kn}$. *Thus* $\mathcal{E} = \{e^{2\pi nit} : n \in \mathbb{Z}\}$ *forms a set of mutually orthogonal functions.*

Orthogonal Polynomials

Example 24 *Let* $X = \mathcal{C}_{\mathbb{R}}[0, 2\pi]$ *and define the inner product* $(f, g) = \int_0^{2\pi} f(t)g(t)\, dt.$

1. $(t^2 + t, 1 - t) = \int_0^{2\pi} (t^2 + t)(1 - t)dt = 2\pi^2(1 - 2\pi^2)$
2. $(\cos(kt), \cos(nt)) = \int_0^{2\pi} \cos(kt)\cos(nt)dt = \frac{\pi}{2}\,\delta_{kn}$. *So* $\{\cos(nt) : n = 0..\infty\}$ *is a mutually orthogonal set.*
3. *Set* $\cos(t) = x$. *Then* $\cos(nt) = \cos(n \arccos(x))$ *becomes a polynomial in* x. *The inner product becomes*

$$(f, g) = \frac{2}{\pi} \int_{-1}^{+1} f(t)g(t) \frac{1}{\sqrt{1 - t^2}}\, dt$$

Orthogonal Polynomials, II

Example 24

(3.) *Set* $T_n(x) = \cos(n \arccos(x))$. *Then* $(T_k, T_n) = \delta_{kn}$, *so that* $\{T_n, n = 0..\infty\}$ *forms an orthogonal set of polynomials. The first few Chebyshev polynomials are* $T_0(x) = 1$ *and*

$$T_1(x) = x$$

$$T_2(x) = 2x^2 - 1$$

$$T_3(x) = 4x^3 - 3x$$

$$T_4(x) = 8x^4 - 8x^2 + 1$$

$$T_5(x) = 16x^5 - 20x^3 + 5x$$

$$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$$

(Go to TOC)

24. Projections

Definition 36 *Let* $X = X_1 \oplus X_2$ *and let* $x = x_1 + x_2$ *be the unique representation of* $x \in X$ *relative to* $X_1 \oplus X_2$*. Then define the mapping* P *by* $P(x) = x_1$*. We call* P *the* projection on X_1 along X_2.

Theorem 52 *Let* $X = X_1 \oplus X_2$ *and* P *be the projection on* X_1 *along* X_2*. Then*

1. $P \in L(X, X)$ *and* $P \in L(X, X_1)$
2. $\mathfrak{R}(P) = X_1$
3. $\mathfrak{N}(P) = X_2$

Pf. $\checkmark$

Example 25 *Let* $X = \mathbb{R}^2$ *and* $P([x_1, x_2]) = x_2$.

Projections, II

Example 26 *Let* $X = \mathbb{P}^3$ *and* $P(\sum_{i=0}^{3} \alpha_i x^i) = \alpha_0 + \alpha_2 x^2$.

Theorem 53 *Let* $P \in L(X, X)$. *Then* P *is a projection on* $\mathfrak{R}(P)$ *along* $\mathfrak{N}(P)$ *iff* $P^2 = P$.

Pf. ($\Rightarrow$) Suppose that P is the projection on $\mathfrak{R}(P)$ along $\mathfrak{N}(P)$. Then $X = \mathfrak{R}(P) \oplus \mathfrak{N}(P)$. Let $x = x_1 + x_2$. Then $P^2(x) = P(P(x_1 + x_2)) = P(x_1) = x_1$ Hence $P^2 = P$.
($\Leftarrow$) Now suppose that $P^2 = P$. (i) Let $y \in \mathfrak{R}(P)$. Then $\exists x \in X$ so that $P(x) = y$. Whence $P(P(x)) = P(y)$. But $P^2 = P$, so $P(P(x)) = P(x) = y$; i.e. $P(y) = y$. If y is also in $\mathfrak{N}(P)$, then $P(y) = 0$ which implies that $y = 0$. Hence $\mathfrak{R}(P) \cap \mathfrak{N}(P) = \{0\}$. (ii) For $x \in X$, $x = P(x) + (I - P)(x)$. Set $x_1 = P(x)$ and $x_2 = (I - P)(x) = x - x_1$. Thence X is equal to $X_1 \oplus X_2$ with P being the projection on X_1 along X_2.

Projection "Symmetry"

Definition 37 $P \in L(X, X)$ *is* idempotent *iff* $P^2 = P$.

Theorem 54 P *is a projection on* X_1 *along* X_2 *iff* $(I - P)$ *is a projection on* X_2 *along* X_1.

Corollary 55 *If* P *is projection, then* $X = \mathfrak{R}(P) \oplus \mathfrak{N}(P)$

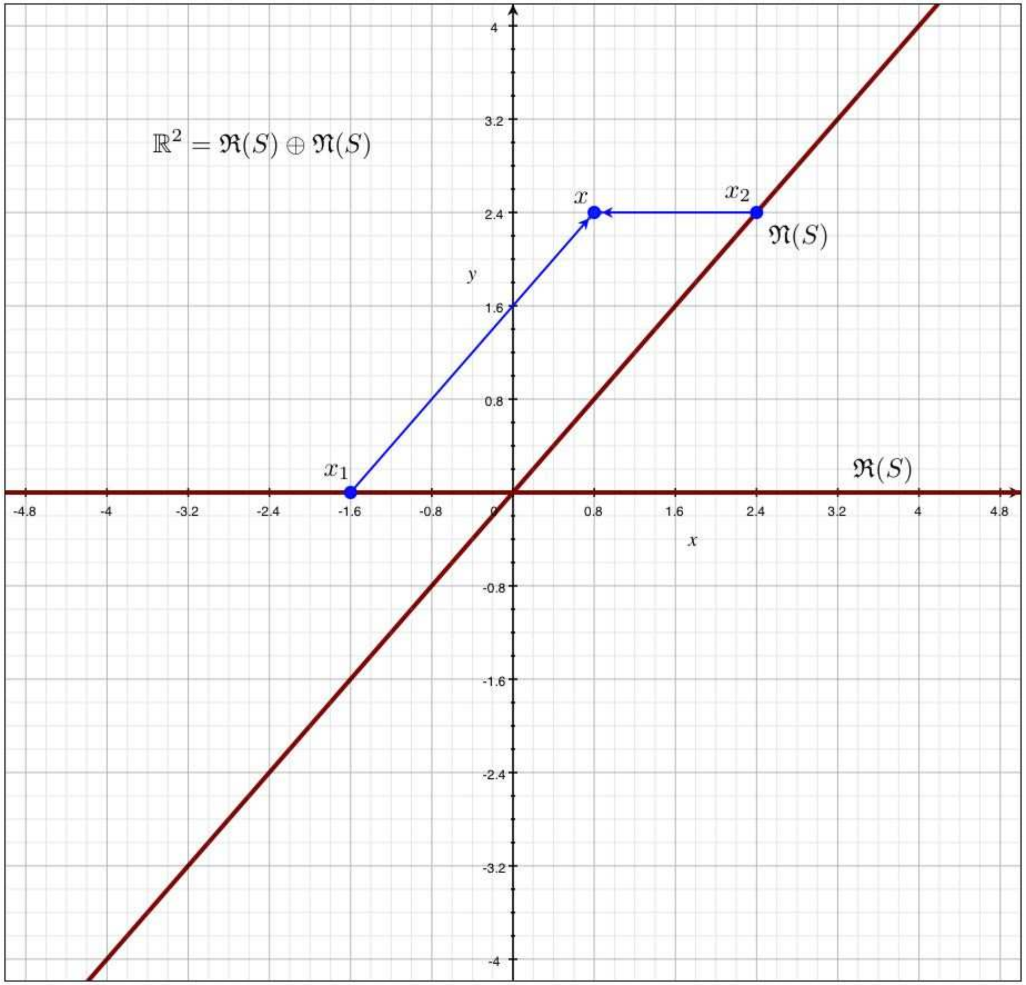
Example Set 27 *Let* $X = \mathbb{R}^2$.

- *Set* $R(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 1 \end{smallmatrix}\right]\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ x_2 \end{smallmatrix}\right]$. *Is* R *a projection?*
- *Set* $S(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) = \left[\begin{smallmatrix} 1 & -1 \\ 0 & 0 \end{smallmatrix}\right]\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} x_1 - x_2 \\ 0 \end{smallmatrix}\right]$. *Is* S *a projection?*
- *Set* $T(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) = \left[\begin{smallmatrix} 0 & 0 \\ 0 & 2 \end{smallmatrix}\right]\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} 0 \\ 2x_2 \end{smallmatrix}\right]$. *Is* T *a projection?*

Definition 38 P *is an* orthogonal projection *on an inner product space iff* $\mathfrak{R}(P) \perp \mathfrak{N}(P)$.

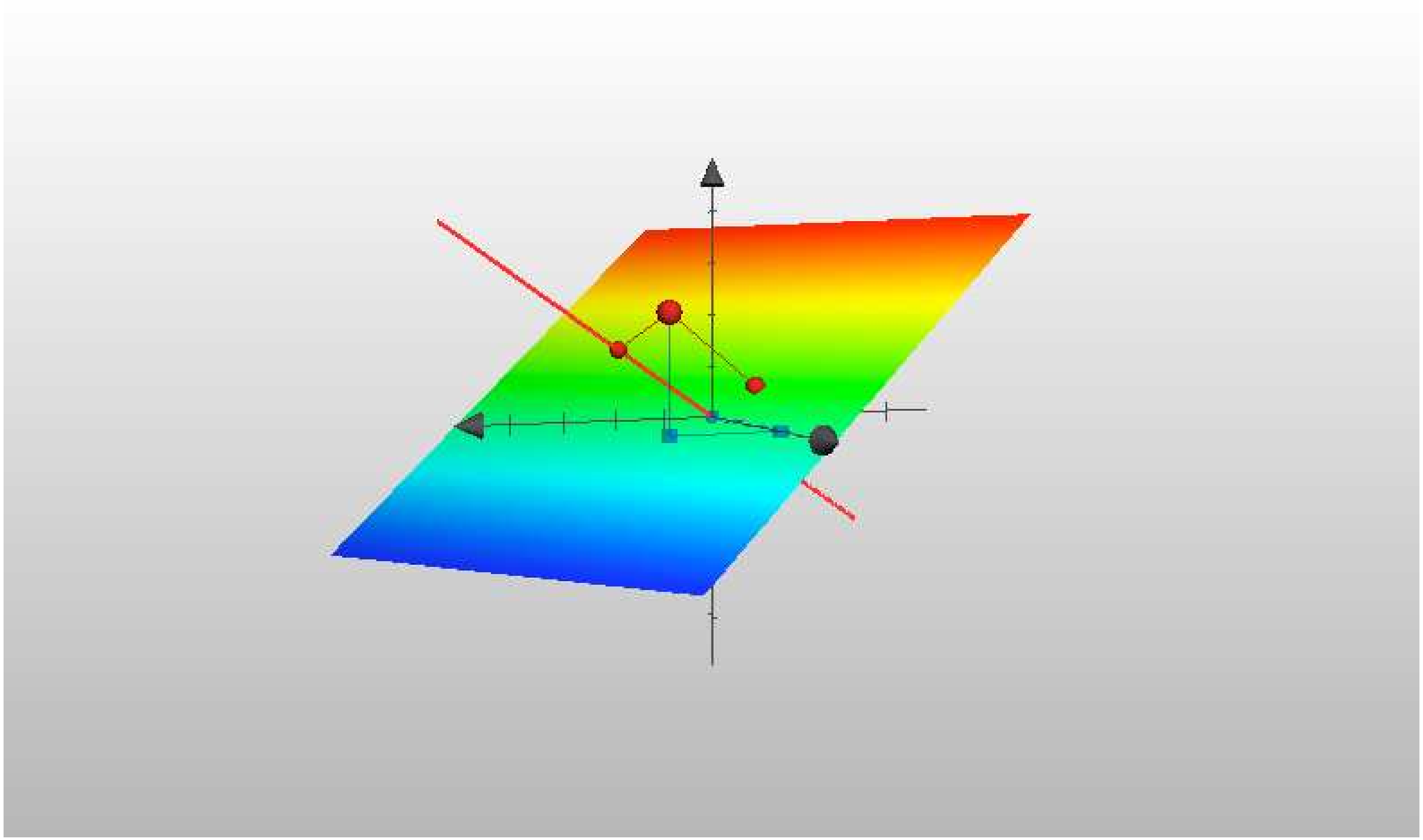
Projection Example S

(Eg.) In $\mathbb{R}^2$, set $S(\left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right]) = \left[\begin{smallmatrix} 1 & -1 \\ 0 & 0 \end{smallmatrix}\right] \left[\begin{smallmatrix} x_1 \\ x_2 \end{smallmatrix}\right] = \left[\begin{smallmatrix} x_1 - x_2 \\ 0 \end{smallmatrix}\right]$.



Projection Example in $\mathbb{R}^3$

(Eg.) In $\mathbb{R}^3$, set $X_1 = \left\{ \begin{bmatrix} -(r+t) \\ r+1 \\ t-1 \end{bmatrix} \right\}$ and $X_2 = \left\{ \begin{bmatrix} 2t \\ 2t \\ t \end{bmatrix} \right\}$.



(Go to TOC)

25. Eigenvalues

Definition 39 *Let* $T \in L(X, X)$, *where* X *is an* n*-dimensional vector space over* F. *A scalar* λ *such that* $T(x) = \lambda x$ *for some nonzero* x *is called an* eigenvalue *of* T.

Theorem 56 *Let* $T \in L(X, X)$ *and let* $\lambda \in F$. *Then* $\mathfrak{N}_\lambda = \{x \mid T(x) = \lambda x\}$ *is a subspace of* X. *This subspace is equal to* $\mathfrak{N}(\lambda I - T)$ *and, if nontrivial, is called an eigenspace.*

Pf. ✓

Background: Let X be an n-dimensional vector space over F. Then $X \cong F^n$. Every linear transformation $T \in L(X, X)$ can be described by its action on a basis. Choosing a basis, allows T to be represented as matrix multiplication (in F). Using the “right” bases gives T as a diagonal matrix greatly simplifying everything.

Eigenvalues, II

- Eigenvalue notes from Luke
- Eigenvalues and differential equations notes from Celes

(Go to TOC)

26. Vector Spaces & Matrices

Background. Let X be a vector space over the field F and let $\dim(X) = n < \infty$.

- Then $X \cong F^n$
- Then, given a basis $\mathcal{B}_X$, each $x \in X$ can be written as $x = [\alpha_1, \ldots, \alpha_n]_{\mathcal{B}_X}$ in "basis order" (row or col format [a])
- Let $T \in L(X, Y)$. T's action on $\mathcal{B}_X$, i.e., the set $T(\mathcal{B}_X)$, completely determines $T(x)$ for any $x \in X$.
- Let $\mathcal{B}_Y$ be a basis for Y. Then there is a *matrix* $\mathbf{T}$ based on $\mathcal{B}_X$ and $\mathcal{B}_Y$, so that

$$y = T(x) = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \times \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix}$$

[a] "Column" and "row" vectors are artifices to aid the arithmetic.

Examples and $\mathbf{T} = [T]$

Example Set 28

1. *Let* $X = \mathbb{R}^2$. *Then* $[1,2]_{\{[1,0],[0,1]\}} = [-1,1]_{\{[1,1],[2,3]\}}$.
2. *Let* $X = \mathbb{P}^2$ *with the "standard basis"* $\{e_i = t^i\}$.
   $x = [1,2,3] = 1 + 2t + 3t^2$
3. *Let* $D : \mathbb{P}^3 \to \mathbb{P}^3$ *be differentiation. Then, with the standard basis* $\{1, t, t^2, t^3\}$,

$$D\left(\begin{bmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{bmatrix}\right) = \begin{bmatrix}0&1&0&0\\0&0&2&0\\0&0&0&3\\0&0&0&0\end{bmatrix} \cdot \begin{bmatrix}\alpha\\ \beta\\ \gamma\\ \delta\end{bmatrix} = \begin{bmatrix}\beta\\ 2\gamma\\ 3\delta\\ 0\end{bmatrix}$$

Definition 40 (The matrix of T) *If* $\mathcal{B}_X = \{e_j\}$ *and* $\mathcal{B}_Y = \{f_i\}$, *then* $\mathbf{T} = [a_{ij}]$ *where* $a_{ij} = \mathrm{proj}_i(T(e_j))$ *with* $\mathrm{proj}_i : Y \to Y$ *being the projection on the* i*th coordinate of* Y *w.r.t.* $\mathcal{B}_Y$.

The Matrix of T

Example Set 29 *Let* $T \in L(\mathbb{R}^2, \mathbb{R}^2)$ *given by* $T([x_1, x_2]) = [x_1 - x_2, x_1 + x_2]$.

1. *Use the standard basis for both.Then* $\mathbf{T} = \left[\begin{smallmatrix} +1 & -1 \\ +1 & +1 \end{smallmatrix}\right]$.
2. *Use* $\mathcal{B}_X = \{[1, 1], [2, 3]\}$ *and* $\mathcal{B}_Y = \{[1, 2], [4, 3]\}$. *Then* $\mathbf{T} = \begin{bmatrix} \frac{12}{5} & -\frac{2}{5} \\ \frac{37}{5} & -\frac{2}{5} \end{bmatrix}$.

—We now return you to the regularly scheduled program.—

Definition 41 (Coordinate representation) *Let* $x \in X$ *and let* $\mathcal{B} = \{e_1, \ldots, e_n\}$ *be a basis for* X. *Then there are unique scalars* ξ_j *such that* $x = \sum_j \xi_j$. *Write* x *in* coordinate representation *with respect to the basis* $\mathcal{B}$ *as* $x = \begin{bmatrix} \xi_1 \\ \vdots \\ \xi_n \end{bmatrix}_{\mathcal{B}}$.

The Transition Matrix

Example 30 *Let* $\mathcal{B} = \{e_1, \ldots, e_4\}$ *be the standard basis for* $\mathbb{R}^4$. *Set* $\mathcal{B}^* = \{[1,2,1,0],[3,3,3,0],[2,-10,0,0],[-2,1,-6,2]\}$. *Then, for* $x \in \mathbb{R}^4$, *define* $T_{\mathcal{B}^* \to \mathcal{B}}$ *by* $\left[\begin{smallmatrix} e_1 & \ldots & e_4 \end{smallmatrix}\right]$ *so*

$$[x]_{\mathcal{B}} = \begin{bmatrix} 1 & 3 & 2 & -2 \\ 2 & 3 & -10 & 1 \\ 1 & 3 & 0 & -6 \\ 0 & 0 & 0 & 2 \end{bmatrix} \times [x]_{\mathcal{B}^*}$$

Hence

$$[x]_{\mathcal{B}^*} = \begin{bmatrix} 1 & 3 & 2 & -2 \\ 2 & 3 & -10 & 1 \\ 1 & 3 & 0 & -6 \\ 0 & 0 & 0 & 2 \end{bmatrix}^{-1} \times [x]_{\mathcal{B}} = \begin{bmatrix} 5 & 1 & -6 & -\frac{27}{2} \\ -\frac{5}{3} & -\frac{1}{3} & \frac{7}{3} & \frac{11}{2} \\ \frac{1}{2} & 0 & -\frac{1}{2} & -1 \\ 0 & 0 & 0 & \frac{1}{2} \end{bmatrix} \times [x]_{\mathcal{B}}$$

Query: How is a vector in $\mathcal{B}_1$ coordinates expressed in $\mathcal{B}_2$ coordinates? Can of cake: use

$$T_{\mathcal{B}_1 \to \mathcal{B}_2} = T^{-1}_{\mathcal{B}_2 \to \mathcal{B}} \times T_{\mathcal{B}_1 \to \mathcal{B}}$$

(Go to TOC)

27. Rank of a Matrix

Theorem 57 *Let* $T \in L(X, Y)$ *where* $\dim(X) = n$ *and* $\dim(Y) = m$. *The* $\rho(T) = r$ *iff there are bases* $\mathcal{B}_X$ *and* $\mathcal{B}_Y$ *such that*

$$\mathbf{T} = \underbrace{\left.\begin{bmatrix} \overbrace{\begin{matrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \end{matrix}}^{r} & \begin{matrix} 0 & \dots & 0 \\ 0 & & \end{matrix} \\ \cdot\;\cdot\;\cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\ \begin{matrix} 0 & \dots & 0 & 1 \end{matrix} & \begin{matrix} 0 & \dots & 0 \end{matrix} \\ \begin{matrix} 0 & \dots & 0 \\ 0 & & \end{matrix} & \begin{matrix} 0 & \dots & 0 \\ 0 & & \end{matrix} \\ \cdot\;\cdot\;\cdot & \cdot\;\cdot\;\cdot \\ \begin{matrix} 0 & \dots & 0 \end{matrix} & \begin{matrix} 0 & \dots & 0 \end{matrix} \end{bmatrix}\right\}}_{n=\dim(X)} m = \dim(Y)$$

The Rank Theorem Examples

Example Set 31

1. *Consider* $T \in L(\mathbb{R}^3 \to \mathbb{R}^2)$. *Then* $\mathbf{T}$ *must have one of the forms below (assuming proper choice of bases):*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

   Explain why.

2. *Consider* $T \in L(\mathbb{R}^3 \to \mathbb{R}^4)$. *Then* $\mathbf{T}$ *must have one of the forms below (assuming proper choice of bases):*

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

   Explain why.

The Rank Theorem Proof

Theorem 57 *Let* $T \in L(X,Y)$. *Then* $\rho(T) = r$ *iff* $\mathbf{T} = \left[\begin{smallmatrix} Id & 0 \\ 0 & 0 \end{smallmatrix}\right]$.

Pf. $(\Leftrightarrow)$ Let $r = \rho(T)$. Choose a basis for $\mathfrak{N}(T)$ of $n-r$ vectors listing it as $\{e_{r+1}, e_{r+2}, \ldots, e_n\}$. Extend this basis to all of X as $\mathcal{B}_X = \{e_1, e_2, \ldots, e_r, e_{r+1}, \ldots, e_n\}$. Calculate $\mathcal{F} = \{T(e_i) \,|\, i = 1..r\}$ which forms a basis for $\mathfrak{R}(T)$. (Thm 3.4.25) Extend $\mathcal{F}$ to a basis $\mathcal{B}_Y$ by adding vectors $\{f_{r+1}, \ldots, f_m\}$. (Thm 3.3.44) Then

$$
\begin{gathered}
f_1 = \mathbf{T}e_1 = (1)f_1 + (0)f_2 + \cdots + (0)f_r + (0)f_{r+1} + \cdots + (0)f_m \\
f_2 = \mathbf{T}e_2 = (0)f_1 + (1)f_2 + \cdots + (0)f_r + (0)f_{r+1} + \cdots + (0)f_m \\
\cdots \\
f_r = \mathbf{T}e_r = (0)f_1 + (0)f_2 + \cdots + (1)f_r + (0)f_{r+1} + \cdots + (0)f_m \\
0 = \mathbf{T}e_{r+1} = (0)f_1 + (0)f_2 + \cdots + (0)f_r + (0)f_{r+1} + \cdots + (0)f_m \\
\cdots \\
0 = \mathbf{T}e_n = (0)f_1 + (0)f_2 + \cdots + (0)f_r + (0)f_{r+1} + \cdots + (0)f_m
\end{gathered}
$$

(Go to TOC)

28. Rank & Algebra of Matrices

Definition 42 *Let* $\mathbf{A} \in \mathfrak{M}_{mn}$ *be the matrix of* $A \in L(X, Y)$ *w.r.t. the bases* $\mathcal{B}_X$ *and* $\mathcal{B}_Y$. *The* rank *of* $\mathbf{A}$ *is the largest number of linearly independent columns in* $\mathbf{A}$.

Theorem 58 *Let* A, B, *and* C *be comparable/conformal matrices and let* $\alpha, \beta \in F$. *Then*

1. $(\mathbf{A}+\mathbf{B})\mathbf{C} = \mathbf{AC} + \mathbf{BC}$
2. $\mathbf{A}(\mathbf{B}+\mathbf{C}) = \mathbf{AB} + \mathbf{AC}$
3. $\mathbf{A}(\mathbf{BC}) = (\mathbf{AB})\mathbf{C}$
4. $(\alpha+\beta)\mathbf{A} = \alpha\mathbf{A} + \beta\mathbf{A}$
5. $\alpha(\mathbf{A}+\mathbf{B}) = \alpha\mathbf{A} + \alpha\mathbf{B}$
6. $(\alpha\mathbf{A})(\beta\mathbf{B}) = (\alpha\beta)(\mathbf{AB})$
7. $\mathbf{A}+\mathbf{B} = \mathbf{B}+\mathbf{A}$
8. $(\mathbf{A}+\mathbf{B})+\mathbf{C} = \mathbf{A}+(\mathbf{B}+\mathbf{C})$

Algebra of Matrices

Theorem 59

1. *The zero matrix* $\mathbf{0} = [0_{ij}]$ *represents the zero transform* $0(x) = 0$ *for every basis.*
2. *The identity matrix* $\mathbf{I} = [\delta_{ij}]$ *represents the identity transform* $I(x) = x$ *for every basis.*
3. *The matrix* $\mathbf{A}$ *is nonsingular iff the transform* A *is nonsingular.*
4. *If* $\mathbf{A}$ *is nonsingular, then* $\mathbf{A}^{-1}$ *is unique.*
5. *If* $\mathbf{A}_n$ *and* $\mathbf{B}_n$ *are nonsingular, then* $(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}$.
6. $\mathrm{rank}(\mathbf{A}_n) = n$ *if and only if* $(\mathbf{A}_n x = 0 \Leftrightarrow x = 0)$.
7. *For* $\mathbf{A} \in \mathfrak{M}_n$, *set* $\mathbf{A}^m = \underbrace{\mathbf{A} \cdot \mathbf{A} \cdot \cdots \cdot \mathbf{A}}_{m}$ *&* $\mathbf{A}^{-m} = (\mathbf{A}^{-1})^m$.

Partitioned Vectors & Matrices

Partitioning a vector or matrix can be very useful and is natural in direct sums. E.g.,

$$\left[\begin{array}{c} x_1 \\ x_2 \\ \hline x_3 \end{array}\right], \quad \left[\begin{array}{cc|c} a_{11} & a_{12} & b_{11} \\ a_{21} & a_{22} & b_{21} \\ \hline c_{11} & c_{12} & d_{11} \end{array}\right], \quad \left[\begin{array}{c|c|c} \mathbf{A}_{11} & \mathbf{A}_{12} & \mathbf{A}_{11} \\ \hline \mathbf{A}_{21} & \mathbf{A}_{22} & \mathbf{A}_{21} \end{array}\right]$$

Theorem 60 *Let* $P \in L(X, X)$ *be a projection and* $\dim(X) = n$. *Then there is a basis for* $X = \mathfrak{R}(P) \oplus \mathfrak{N}(P)$ *s.t.*

$$\mathbf{P} = \left[\begin{array}{cccc|ccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & \cdot \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \hline 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdots & \cdot & \cdot & \cdots & . \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{array}\right]$$

(Go to TOC)

29. Similarity & Equivalence

Ab hinc: X and Y are vector spaces over F with $\dim X = n$ and $\dim(Y) = m$.

Theorem 61 *Let* $\mathcal{B}_X = \{e_1, \ldots, e_n\}$ *be a basis for* X *and let* $\mathrm{P} = [p_{ij}]$ *be an* $n \times n$ *matrix. Set* $e'_k = \sum_j p_{jk} e_j$. *Then* $\mathcal{B}'_X = \{e'_1, \ldots, e'_n\}$ *is a basis for* X *iff* P *is nonsingular.*

Pf. *Calculation based on the linear independence of* $\mathcal{B}_X$.

Definition 43 *Let* P *be the matrix of Thm 61, then* P *is* the matrix of $\mathcal{B}'_X$ w.r.t $\mathcal{B}_X$.

Theorem 62 P *is the matrix of* $\mathcal{B}'_X$ *w.r.t* $\mathcal{B}_X$ *iff* P^{-1} *is the matrix of* $\mathcal{B}_X$ w.r.t $\mathcal{B}'_X$.

Pf. *Exercise*.

Similarity of Matrices

Theorem 63 *Let* P *be the matrix of* $\mathcal{B}'_X$ *w.r.t* $\mathcal{B}_X$ *and* Q *be the matrix of* $\mathcal{B}''_X$ *w.r.t* $\mathcal{B}'_X$*. Then* PQ *is the matrix of* $\mathcal{B}''_X$ *w.r.t* $\mathcal{B}_X$*.*

Pf. *Exercise*.

Theorem 64 *Let* P *be the matrix of* $\mathcal{B}'_X$ *w.r.t* $\mathcal{B}_X$ *and let* $x \in X$ *be* $\mathbf{x}$ *in* $\mathcal{B}_X$ *coordinates. Then* $\mathrm{P}\mathbf{x}' = \mathbf{x}$ *gives* x *in* $\mathcal{B}'_X$ *coordinates.*

Pf. *Exercise*.

$$
\begin{array}{ccc}
\mathbf{x} = \mathbf{P}\mathbf{x}' & \xrightarrow{\mathbf{A}} & \mathbf{y} = \mathbf{A}\mathbf{x} \\
\Big\uparrow \mathbf{P} & & \Big\downarrow \mathbf{Q} \\
\mathbf{x}' & \xrightarrow{\mathbf{A}'} & \begin{aligned} \mathbf{y}' &= \mathbf{Q}\mathbf{y} \\ &= \mathbf{QAP}\mathbf{x}' \end{aligned}
\end{array}
$$

(Go to TOC)

30. Equivalence of Transformations

Theorem 65 *Let* $A \in L(X, Y)$ *where*

- A *has matrices* $\mathbf{A}_{\mathcal{B}_X \to \mathcal{B}_Y}$, *and* $\mathbf{A}'_{\mathcal{B}'_X \to \mathcal{B}'_Y}$, *resp.*
- $\mathbf{P}$ *is the matrix of* $\mathcal{B}'_X$ *w.r.t.* $\mathcal{B}_X$ *and* $\mathbf{Q}$ *of* $\mathcal{B}'_Y$ *w.r.t.* $\mathcal{B}_Y$

Then $\mathbf{A}' = \mathbf{QAP}$.

Pf.

$$Ae'_i = A \cdot \sum_k p_{ki} e_k = \sum_k p_{ki} A e_k = \sum_k p_{ki} \left(\sum_l a_{lk} f_l \right)$$

$$= \sum_k p_{ki} \left(\sum_l a_{lk} \left[\sum_j q_{jl} f'_j \right] \right) = \sum_k \sum_l \sum_j q_{jl} a_{lk} p_{ki} \cdot f'_j$$

Whence

$$a'_{ij} = \sum_l \sum_k q_{il} a_{lk} p_{kj}$$

Definition of Equivalence

Definition 44 *Two* $m \times n$ *matrices* $\mathbf{A}$ *and* $\mathbf{A}'$ *are* equivalent *iff there are nonsingular square matrices* $\mathbf{P}_n$ *and* $\mathbf{Q}_m$ *such that* $\mathbf{A}' = \mathbf{Q_m} \cdot \mathbf{A} \cdot \mathbf{P_n}$. *Equivalence is written as* $\mathbf{A}' \sim \mathbf{A}$.

Theorem 66 *Matrix equivalence is an equivalence relation. I.e.,* $\sim$ *is reflexive, symmetric, and transitive.*

Pf. *Exercise*.

Theorem 67 *Let* $\mathbf{A}$ *and* $\mathbf{B} \in \mathfrak{M}_{m,n}$. *Then*

1. $\mathbf{A}$ *is equivalent to* $\left[\begin{array}{c:c} I_r & 0 \\ \hdashline 0 & 0 \end{array}\right]$ *where* $r = \mathrm{rank}(A)$.
2. $A \sim B$ *iff* $\mathrm{rank}(A) = \mathrm{rank}(B)$.

Equivalence Example

Example 32 *Consider* $A \in L(\mathbb{R}^4, \mathbb{R}^5)$.

Suppose $\mathbf{A} = \begin{bmatrix} 7 & -9 & 5 & -4 \\ 7 & 3 & -8 & -5 \\ 4 & 9 & 5 & 6 \\ 11 & 0 & 10 & 2 \\ 0 & 12 & -13 & -1 \end{bmatrix}$ *and* $\mathbf{A}' = \begin{bmatrix} 1 & 13 & 0 & 12 \\ -21 & -31 & 8 & 6 \\ 13 & 14 & -7 & 15 \\ -1 & 21 & 3 & 0 \\ 11 & -46 & -10 & -21 \end{bmatrix}$.

Then $\mathbf{P} = \begin{bmatrix} -1 & 1 & 1 & 0 \\ 1 & 0 & -1 & -1 \\ 1 & 1 & -1 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$ and $\mathbf{Q} = \begin{bmatrix} -1 & 0 & -1 & 1 & -1 \\ 0 & 1 & -1 & -1 & 0 \\ 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & 0 & 1 & 1 \\ -1 & 0 & 0 & -1 & 1 \end{bmatrix}$.

1. Show that $\mathbf{A}' = \mathbf{QAP}$.
2. Find the matrix $\left[\begin{array}{c|c} I_r & 0 \\ \hline 0 & 0 \end{array}\right]$ equivalent to both $\mathbf{A}$ and $\mathbf{A}'$.

(Go to TOC)

31. Determinants: “Work Day”

Exercise. Use an undergraduate text on Linear Algebra to:

1. Define the *determinant of a matrix*.
2. List the main properties/theorems on determinants.
3. Choose the five most important properties.
4. Give a numerical example demonstrating each important property listed.
5. Discuss the relation between *determinant of a matrix* and *nonsingularity of a linear transformation*.

(Go to TOC)

32. Determinants & Invariants

Recall the following

Theorem 68 *Let* $A \in L(X, X)$.

- $|A| \neq 0$ *iff* A *is nonsingular*
- $|A \cdot B| = |A| \cdot |B|$
- $|A^{-1}| = |A|^{-1}$
- $|\alpha A| = \alpha^n |A|$ *where* $n = dim(A)$
- $|A| = 0$ *iff*
  - A *has a row/column of zeros*
  - A *has two identical rows/columns*
  - A *has a row/column that is a linear combination of other rows/columns*
  - $A\mathbf{x} = \mathrm{b}$ *has nonunique solutions*

Eigenvalue & Eigenvector

Definition 45 *Let $A \in L(X, X)$. A scalar λ such that there is a nonzero $x \in X$ for which $Ax = \lambda x$ is an* eigenvalue *and the corresponding x is an* eigenvector*.*

Definition 46 *The polynomial $p(\lambda) = |A - \lambda I|$ is the* characteristic polynomial of A*.*

Theorem 69 (Cayley-Hamilton) *Let $A \in L(X, X)$. Then $p(A) = 0$. (NB: Also $p(0) = \prod \lambda_i$. See Zhou*[a] *)*

Definition 47 *Let $A \in L(X, X)$. Then the subspace Y is an* invariant subspace under A *iff $A(Y) \subseteq Y$; i.e., $\forall y \in Y$, we have $Ay \in Y$.*

Definition 48 *Set $\mathfrak{N}_\lambda(A) = \mathfrak{N}_\lambda = \mathfrak{N}(A - \lambda I)$.*

---

[a] "Intro. to Symmetric Polynomials & Symmetric Functions"

Reduced Linear Transformation

Theorem 70 *Let* $A \in L(X,X)$*. Then* X, $\mathfrak{R}(A)$, $\mathfrak{N}(A)$, *and* $\{0\}$ *are all invariant subspaces under* A.

Example 33 *Let* $A : \mathbb{R}^2 \to \mathbb{R}^2$ *be given by* $\left[\begin{smallmatrix} 1 & 2 \\ 2 & 4 \end{smallmatrix}\right]$. *Then* $\{0\}$. $\mathfrak{N}(A) = \langle \left[\begin{smallmatrix} 2 \\ -1 \end{smallmatrix}\right] \rangle$, $\mathfrak{R}(A) = \langle \left[\begin{smallmatrix} 1 \\ 2 \end{smallmatrix}\right] \rangle$, *and* $\mathbb{R}^2$ *are all invariant.*

Example 34 *(Exercise.) Let* $B : \mathbb{R}^3 \to \mathbb{R}^3$ *be given by* $\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 1 \\ 1 & 2 & 0 \end{bmatrix}$. *Then* $\{0\}$, $\mathfrak{N}(B) = \langle$? $\rangle$, $\mathfrak{R}(B) = \langle ? \rangle$, *and* $\mathbb{R}^3$ *are all invariant.*

Theorem 71 *Let* λ *be an eigenvalue of* $A \in L(X,X)$*. Then* $\mathfrak{N}_\lambda$ *is invariant. (Exercise.)*

Definition 49 *Let* $X = Y \oplus Z$ *be such that both* Y *and* Z *are invariant subspaces under* $A \in L(X,X)$*. Then* A *is* reduced *by* Y *and* Z *and* A *can take form* $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$.

(Go to TOC)

33. Eigenvalues & Diagonalization

Ab hinc: X is an n-dimensional vector space over F.

Theorem 72 *Let* $\{\lambda_i \,|\, i = 1..p\}$ *be a set of distinct eigenvalues of* $A \in L(X, X)$ *with corresponding nonzero eigenvectors* $\mathcal{E} = \{e_i' \,|\, i = 1..p\}$. *Then* $\mathcal{E}$ *is linearly independent.*

Pf. Assume $\mathcal{E}$ is dependent. Choose the smallest set of vectors from $\mathcal{E}$ such that $0 = \sum_{i=1}^{r} \alpha_i e_i'$ (reordering the $r \leq p$ vectors as needed). Then $0 = A(0) = A\left(\sum_{i=1}^{r} \alpha_i e_i'\right)$ which gives $0 = \sum_{i=1}^{r} (\lambda_i \alpha_i e_i')\ ^{(*)}$. Now $0 = \lambda_r 0 = \lambda_r \sum_{i=1}^{r} \alpha_i e_i'$, or $0 = \sum_{i=1}^{r} \lambda_r \alpha_i e_i'\ ^{(**)}$. Subtract $^{(*)}$ from $^{(**)}$ to obtain $0 = \sum_{i=1}^{r-1} (\lambda_r - \lambda_i) \alpha_i e_i'$ which contradicts r being minimal. Hence $\mathcal{E}$ is linearly independent.

“Eigenbasis”

Theorem 73 *If $A \in L(X,X)$ has n distinct eigenvalues, then there is a basis of eigenvectors $\mathcal{B}_e = \{e_i' \mid i = 1..p\}$ such that the matrix of A is $\mathrm{diag}(\lambda_1, \ldots, \lambda_n)$.*

Pf. *Exercise.*

Corollary 74 *If $A \in L(X,X)$ has n distinct eigenvalues, then every matrix for A is similar to a diagonal matrix.*

Pf. Collect the eigenvectors $\mathcal{E} = \{e_i' \mid i = 1..n\}$. Set $P = [e_1', \ldots, e_n']$. Then $\mathrm{diag}(\lambda_1, \ldots, \lambda_n) = P^{-1} A\, P$.

Example 35 *See the Maple worksheet. (To see what happens without a “cooked” example, enter the following in Maple:* `with(LinearAlgebra):` `A := RandomMatrix(5,5, generator=rand(-3..3)); Eigenvectors(A);`*)*

(Go to TOC)

34. "Eigen-Basis" Examples

Example 36

1. *Let* $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \in L(\mathbb{R}^3, \mathbb{R}^3)$. *Then* $p(\lambda) = \lambda^3 - 2\lambda^2 + \lambda$ *(and* $m(\lambda) = \lambda^2 - \lambda$*) which indicates that* A *has eigenvalues:* $0, 1, 1$. *The corresponding eigenvectors come from* $\mathfrak{N}_\lambda = \mathfrak{N}(A - \lambda I)$. *So* $\mathfrak{N}_0 = \langle \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} \rangle$ *and* $\mathfrak{N}_1 = \langle \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \rangle$. *(Found by solving* $\begin{bmatrix} 1-\lambda & 0 & 0 \\ 1 & -\lambda & 0 \\ 0 & 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$ *with* $\lambda = 0$ *and* 1, *respectively.)*
   - (a) *Define* P *and find* P^{-1}.
   - (b) *Calculate the diagonal matrix* $P^{-1} A P$ without *using matrix multiplication.*

“Eigen-Basis” Examples, II

Example 37

1. *Let* $B = \begin{bmatrix} 1 & 1 & 2 \\ 1 & 1 & 2 \\ 1 & 0 & 1 \end{bmatrix} \in L(\mathbb{R}^3, \mathbb{R}^3)$. *Then* $p(\lambda) = \lambda^3 - 3\lambda^2$ *(and* $m(\lambda) = \lambda^3 - 3\lambda^2$*) which indicates that* B *has eigenvalues:* $0, 0,$ *and* 3. *The corresponding eigenvectors come from* $\mathfrak{N}_\lambda = \mathfrak{N}(B - \lambda I)$. *So* $\mathfrak{N}_0 = \langle \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \rangle$ *and* $\mathfrak{N}_3 = \langle \begin{bmatrix} 2 \\ 2 \\ 1 \end{bmatrix} \rangle$. *(Found by solving* $\begin{bmatrix} 1-\lambda & 1 & 2 \\ 1 & 1-\lambda & 2 \\ 1 & 0 & 1-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = 0$ *with* $\lambda = 0,$ *and* 3, *respectively.)*
   - (a) *Explain why* P *(and so* P^{-1}*) doesn't exist.*
   - (b) *Can* B *be diagonalized? Why or why not?*

(Solution.)

(Go to TOC)

35. Geometric Multiplicity

Definition 50 *Let λ be an eigenvalue of $A \in L(X, X)$. Then*

- *the* algebraic multiplicity of λ *is the multiplicity as a root of the characteristic polynomial* $p(\lambda)$;
- *the* geometric multiplicity of λ *is the dimension of the nullspace* $\mathfrak{N}_\lambda = \mathfrak{N}(A - \lambda I)$.

Example 38 *Let $X = \mathbb{R}^3$. Each matrix below has characteristic polynomial* $p(\lambda) = -(\lambda - 2)^3$.

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

alg $= 3$, *geo* $= 3$ *alg* $= 3$, *geo* $= 2$ *alg* $= 3$, *geo* $= 1$

$iev = \{e_1, e_2, e_3\}$ $iev = \{e_1, e_3\}$ $iev = \{e_1\}$

Reduction Partition

Theorem 75 *Let* $X = X_1 \oplus X_2$ *be a direct sum that reduces* $A \in L(X, X)$; *i.e.,* A *is invariant on* X_1 *and* X_2. *Then there is a basis* $\mathcal{B}$ *for* X *such that*

$$A_{\mathcal{B}} = \left[\begin{array}{c:c} A_1 & 0 \\ \hdashline 0 & A_2 \end{array}\right]$$

Theorem 76 *Let* $X = X_1 \oplus \cdots \oplus X_p$ *be a direct sum that reduces* $A \in L(X, X)$; *i.e.,* $A_k = A|_{X_k}$ *is invariant on* X_k *for* $k = 1..p$. *Then there is a basis* $\mathcal{B}$ *for* X *such that*

$$A_{\mathcal{B}} = \begin{bmatrix} A_1 & 0 & \ldots & 0 \\ 0 & A_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & A_p \end{bmatrix} \quad \textit{and} \quad |A_{\mathcal{B}}| = \prod_{k=1}^{p} |A_k|$$

Minimal Polynomial

Example 39 *If* $A \in L(X,X)$ *has* n *distinct eigenvalues in* F*, then* $X = \mathfrak{N}_{\lambda_1} \oplus \cdots \oplus \mathfrak{N}_{\lambda_n}$ *and* $A = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$.

Definition 51 *Let* $A \in L(X,X)$. *Then there is a monic polynomial* $m(\lambda)$, *the* minimal polynomial, *such that*

- $m(A) = 0$
- *any polynomial* m' *with* $m'(A) = 0$ *has* $\deg(m) \leq \deg(m')$

Example 40 *The three matrices of Example 38 have*

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \qquad \begin{bmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$$

$$m(\lambda) = (\lambda - 2) \qquad m(\lambda) = (\lambda - 2)^2 \qquad m(\lambda) = (\lambda - 2)^3$$

Properties of the Minimal Polynomial

Theorem 77 *The minimal polynomial* $m(\lambda)$ *is unique.*

Theorem 78 *Let* $q(\lambda)$ *be a polynomial such that* $q(A) = 0$. *Then* $m(\lambda) | \, q(\lambda)$.

Corollary 79 *The minimal polynomial divides the characteristic polynomial; i.e.,* $m(\lambda) | \, p(\lambda)$.

Theorem 80 *The characteristic polynomial divides a power of the minimal polynomial:* $p(\lambda) \,\big|\, [m(\lambda)]^n$ *where* $n = \dim(X)$.

Corollary 81 $m(\lambda) \,\big|\, p(\lambda) \,\big|\, [m(\lambda)]^n$.

Proofs. *Exercises*.

(Go to TOC)

36. Jordan Canonical Form

Definition 52 *Let* $A \in L(X, X)$. *If there is a power* k *such that* $A^k = 0$, *but* $A^{k-1} \neq 0$, *then* A *is* nilpotent of index k.

NB: For a nilpotent matrix A of index k, $m_A(\lambda) = \lambda^k$.

Definition 53 *Define* $N_k \in \mathfrak{M}_{k \times k}$ *to be*

$$N_k = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}$$

Then N_k *is nilpotent of index* k.

Definition 54 *Define* $J_k(\lambda_i) = N_k + \lambda_i I$. *The matrix* $J_k(\lambda_i)$ *is a* Jordan block.

Jordan Blocks

Theorem 82 *Let* $J_k(\lambda_i)$ *be the Jordan block matrix of size* k *(with eigenvalue* λ_i*). Then*

- $p(\lambda) = (\lambda - \lambda_i)^k$
- $m(\lambda) = (\lambda - \lambda_i)^k$

Pf. Consider $J_k(\lambda_i) - \lambda I$. For $\lambda = \lambda_i$, we see

$$J_k(\lambda_i) - \lambda_i I = N_k$$

which is nilpotent of index k. Hence $p(\lambda) = (\lambda - \lambda_i)^k$. Since $m \,|\, p$ and no lower power of N_k than k gives 0, it follows that we also have $m(\lambda) = (\lambda - \lambda_i)^k$.

Jordan Form

Theorem 83 *Let $A \in L(X, X)$ with*

$$p(\lambda) = \prod_{i=1}^{r} (\lambda - \lambda_i)^{n_i} \quad \text{and} \quad m(\lambda) = \prod_{i=1}^{r} (\lambda - \lambda_i)^{m_i}$$

Then there is a block-diagonal matrix for A with blocks $J(\lambda_i)$. For each λ_i the blocks $J(\lambda_i)$ have the properties:

1. *There is at least one block $J_{m_i}(\lambda_i)$; all others have order $\leq m_i$.*
2. *The sum of the orders of the blocks for $J(\lambda_i)$ is n_i.*
3. *The number of blocks $J(\lambda_i)$ equals the geometric multiplicity of λ_i.*
4. *A uniquely determines the number of blocks $J(\lambda_i)$.*

JCF Example, I

Example 41 *Let* $A = \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 3 & 1 & 0 \\ 1 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$. *($A$ is not diagonalizable.)*

First $p(\lambda) = (4-\lambda)(1-\lambda)^3$. *Hence,* $m(\lambda) = (4-\lambda)(1-\lambda)^k$ *where* $k = 1, 2,$ *or* 3. *Testing, we determine that* $m(\lambda) = (4-\lambda)(1-\lambda)^2$. *Thus there are Jordan blocks* $J_1(4)$ *and* $J_2(1)$ *from the factor powers in* $m(\lambda)$. *Since the sum of the block's indices must be* 4, *the last block is* $J_1(2)$. *We have determined that*

$$JCF(A) = \mathrm{diag}(J_1(4), J_2(1), J_1(1)) = \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Now we compute the transition matrix P *that converts* $JCF(A) = P^{-1}AP$. *The first eigenvalue* 4 *has one independent eigenvector:* $\mathfrak{N}_4 = \langle [1, 1, 1, 0]^{\mathrm{T}} \rangle$. *The second*

JCF Example, II

Example 41 *(continued) eigenvalue* 1 *has geometric multiplicity* 2*; i.e.,* $\dim(\mathfrak{N}_1) = 2$, *and has only* 2 *independent eigenvectors.* $\mathfrak{N}_1 = \langle [0,0,0,1]^{\mathrm{T}}, [1,1,-2,0]^{\mathrm{T}} \rangle$. *We need another independent vector for* P. *Set* $\mathfrak{N}_{1,2} = \mathfrak{N}\left((A - 1I)^2\right)$ $= \langle [0,0,0,1]^{\mathrm{T}}, [-3,0,1,0]^{\mathrm{T}}, [-5,1,0,0]^{\mathrm{T}} \rangle$. *Let* $N_\lambda = A - \lambda I$, *then* x_1, $x_2 = N_1 x_1, \ldots, x_j = N_1 x_{j-1}$ *can form an independent* chain *of vectors. Try* $x_1 = [-5,1,0,0]^{\mathrm{T}}$, *then* $x_2 = N_1 x_1 = [2,2,-4,0]^{\mathrm{T}}$ *and* $x_3 = N_1 x_2 = 0$. *We have* 4 *independent vectors with which to construct* P.

$$P = [\mathfrak{N}_4[1], \mathfrak{N}_1[1], x_2, x_1] = \begin{bmatrix} 1 & 0 & 2 & -5 \\ 1 & 0 & 2 & 1 \\ 1 & 0 & -4 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \text{ and } \begin{bmatrix} 4 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = P^{-1}AP$$

(Go to TOC)

37. JCF Examples

We will consider a 3×3 matrix A over $\mathbb{R}$; i.e., $A \in L(\mathbb{R}^3, \mathbb{R}^3)$.

Example 42 *Let* $A = \begin{bmatrix} 4 & 4 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}$. *The characteristic polynomial is* $p(\lambda) = (4-\lambda)(2-\lambda)^2$. *The minimal polynomial is found to be the same:* $m(\lambda) = (4-\lambda)(2-\lambda)^2$. *We can directly write the Jordan canonical form* $JCF(A) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. *(Explain why.) We construct the transition matrix* P *from eigen- and generalized eigenvectors. Set* $N_4 = (A - 4I)$. *The nullspace of* N_4 *is* $\langle [1,0,0]^{\mathrm{T}} \rangle$. *Set* $v_1 = [1,0,0]^{\mathrm{T}}$. *Now consider* $N_2 = (A - 2I)$. *The nullspace of* N_2 *is* $\langle [-2,1,0]^{\mathrm{T}} \rangle$. *This space is "too small," since the algebraic multiplicity of* $\lambda = 2$ *is* 2. *Consider* $N_{2,2} = (A - 2I)^2$. *The nullspace of* $N_{2,2}$ *is* $\langle [-2,1,0]^{\mathrm{T}}, [-9/2,0,1]^{\mathrm{T}} \rangle$. *Set* $x_2 = [-9/2,0,1]^{\mathrm{T}}$. *This choice*

JCF Examples, II

Example 42 (continued) *gives* $x_2 \in \mathfrak{N}(N_{2,2}) - \mathfrak{N}(N_2)$, *and then define* $x_1 = N_2\, x_2 = [-8, 4, 0]^{\mathrm{T}}$. *Let* P *be the matrix* $[v_1, x_1, x_2]$. *Then*

$$P = \begin{bmatrix} 1 & -8 & -9/2 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad P^{-1} = \begin{bmatrix} 1 & 2 & 9/2 \\ 0 & 1/4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Whence

$$JCF(A) = P^{-1}\, A\, P = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}.$$

JCF Finite Field Example

We will consider the same 3×3 matrix A, but now over $\mathbb{Z}_5$; i.e., $A \in L(\mathbb{Z}_5^3, \mathbb{Z}_5^3)$. Proceed as before, but calculate in $\mathbb{Z}_5$.

Example 43 *Let* $A = \begin{bmatrix} 4 & 4 & 1 \\ 0 & 2 & 4 \\ 0 & 0 & 2 \end{bmatrix}$. *The characteristic polynomial is* $p(\lambda) = (4-\lambda)(2-\lambda)^2$. *The minimal polynomial is found to be the same:* $m(\lambda) = (4-\lambda)(2-\lambda)^2$. *We can directly write the Jordan canonical form* $JCF(A) = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix}$. *(Explain why.) We again construct the transition matrix* P *from eigen- and generalized eigenvectors. Set* $N_4 = (A - 4I) = (A + 1I) \mod 5$. *The nullspace of* N_4 *is* $\langle [1,0,0]^{\mathrm{T}} \rangle$. *Set* $v_1 = [1,0,0]^{\mathrm{T}}$. *Now consider* $N_2 = (A - 2I) = (A + 3I) \mod 5$. *The null-space of* N_2 *is* $\langle [3,1,0]^{\mathrm{T}} \rangle$. *This space is "too small," since the algebraic multiplicity of* $\lambda = 2$ *is* 2.

JCF Finite Field Example, II

Example 43 (continued) *Consider* $N_{2,2} = (A - 2I)^2$ $= (A + 3I)^2 \mod 5 = \begin{bmatrix} 4 & 3 & 3 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$. *The nullspace of* $N_{2,2}$ *is* $\langle [3, 1, 0]^{\mathrm{T}}, [3, 0, 1]^{\mathrm{T}} \rangle$. *Set* $x_2 = [3, 0, 1]^{\mathrm{T}}$ *as this choice gives* $x_2 \in \mathfrak{N}(N_{2,2}) - \mathfrak{N}(N_2)$, *and then define* $x_1 = N_2\, x_2 = [2, 4, 0]^{\mathrm{T}}$. *Let* P *be the matrix* $[v_1, x_1, x_2]$. *Then*

$$P = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \textit{and} \quad P^{-1} = \begin{bmatrix} 1 & 2 & 2 \\ 0 & 4 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Whence

$$JCF(A) = P^{-1}\, A\, P = \begin{bmatrix} 4 & 40 & 25 \\ 0 & 32 & 16 \\ 0 & 0 & 2 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{bmatrix} \mod 5.$$

JCF Exercise

Exercise. Set $A = \begin{bmatrix} 0 & -1 & 0 & 2 \\ -1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 \\ -1 & 1 & 0 & 1 \end{bmatrix}$.

Find

1. $p(\lambda)$
2. $m(\lambda)$
3. the transition matrix P
4. $JCF(A)$

THE END

(Go to TOC)