Section 2: Linear maps

Matthew Macauley

School of Mathematical & Statistical Sciences
Clemson University
http://www.math.clemson.edu/~macaule/

Math 8530, Advanced Linear Algebra

Preliminaries

Goal

Abstract the concept of a matrix as a linear mapping between vector spaces.

Advantages:

- simple, transparent proofs;
- better handles infinite dimensional spaces.

Definition (revisted)

A linear map (or *mapping*, *transformation*, or *operator*) between vector spaces X and U over K is a function $T\colon X \to U$ that is:

(i) additive: $T(x+y) = T(x) + T(y), \quad \text{for all } x, y \in X,$

(ii) homogeneous: $T(ax) = aT(x), \quad \text{for all } x \in X,\ a \in K.$

The domain space is X and the target space is U.

Usually we'll write Tx for $T(x)$, and so additivity is just the distributive law:

$$T(x+y) = Tx + Ty.$$

Examples of linear maps

(i) Any isomorphism;

(ii) $X = U = \{\text{polynomials of degree } < n \text{ in t}\}, \quad T = \frac{d}{dt}$.

(iii) $X = U = \mathbb{R}^2, \quad T =$ rotation about the origin.

(iv) X any vector space, $U = K$ (1-dimensional), T any $\ell \in X'$.

(v) $X = U = \mathcal{C}([0,1], \mathbb{R})$, $g \in X$. $\quad (Tf)(x) = \displaystyle\int_0^1 f(y)g(x-y)\,dy$.

(vi) $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, $u = Tx$, where $u_i = \displaystyle\sum_{j=1}^{n} t_{ij}x_j, \quad i = 1, \dots, m$.

(vii) $X = U = \{\text{piecewise cont. } [0,\infty) \to \mathbb{R} \text{ of ``exponential order''}\}$,
$(Tf)(s) = \displaystyle\int_0^\infty f(t)e^{-st}\,dt$. "*Laplace transform*"

(viii) $X = U = \{\text{functions with } \int_{-\infty}^{\infty} |f(x)|\,dx < \infty\}$,
$(Tf)(\xi) = \displaystyle\int_{-\infty}^{\infty} f(x)e^{i\xi x}\,dx$. "*Fourier transform*"

Basic properties of linear maps

Theorem 2.1

Let $T\colon X \to U$ be a linear map.

(a) The image of a subspace of X is a subspace of U.

(b) The preimage of a subspace of U is a subspace of X.

(Proof is a HW exercise.) □

Definition

The range of T is the image $R_T := T(X)$. The rank of T is $\dim R_T$.

The nullspace (or "kernel") of T is the preimage of 0:

$$N_T := T^{-1}(0) = \{x \in X \mid Tx = 0\}\,.$$

The nullity of T is $\dim N_T$.

Remark

A linear map $T\colon X \to U$ is 1–1 if and only if $N_T = \{0\}$.

The rank-nullity theorem

Theorem 2.2

Let $T \colon X \to U$ be a linear map. Then $\dim R_T + \dim N_T = \dim X$.

Proof

Consequences of the rank-nullity theorem

Corollary A

Suppose $\dim U < \dim X$. Then $Tx = 0$ for some $x \neq 0$.

Proof

Example A

Take $X = \mathbb{R}^n$, $U = \mathbb{R}^m$, with $m < n$. Let $T \colon \mathbb{R}^n \to \mathbb{R}^m$ be any linear map (see Example (vi)).

Since $m = \dim U < \dim X < n$, Corollary A implies that the system of m equations

$$\sum_{j=1}^{n} t_{ij} x_j = 0 \qquad i = 1, \ldots, m$$

has a non-trivial solution, i.e., not all $x_j = 0$.

Consequences of the rank-nullity theorem

Corollary B

Suppose $\dim X = \dim U < \infty$ and the only vector satisfying $Tx = 0$ is $x = 0$. Then $R_T = U$.

Proof

Example B

Take $X = U = \mathbb{R}^n$, and $T\colon \mathbb{R}^n \to \mathbb{R}^n$ given by $\sum_{j=1}^{n} t_{ij}x_j = u_i$, for $i = 1, \dots, n$.

If the related homogeneous system of equations $\sum_{j=1}^{n} t_{ij}x_j = 0$, for $i = 1, \dots, n$, has only the trivial solution $x_1 = \cdots = x_n = 0$, then the inhomogeneous system T has a unique solution for any choice of $u_1 \dots, u_n$.

[*Reason*: $T\colon \mathbb{R}^n \to \mathbb{R}^n$ is an isomorphism.]

Polynomial interpolation

Let $X = \{p \in \mathbb{C}[x] \mid \deg p < n\}$ and $U = \mathbb{C}^n$.

Pick any distinct $s_1, \dots, s_n \in \mathbb{C}$, and define

$$T \colon X \longrightarrow U, \qquad T \colon p \mapsto \big(p(s_1), \dots, p(s_n)\big).$$

Suppose $Tp = 0$ for some $p \in X$.

Then $p(s_1) = \cdots = p(s_n) = 0$, which is impossible because p has at most $n-1$ distinct roots.

Therefore $N_T = \{0\}$, and so Corollary B implies that $R_T = U$.

Average value of a polynomial

Let $X = \{p \in \mathbb{R}[x] \mid \deg p < n\}$ and $U = \mathbb{R}^n$.

Let $I_1, \dots, I_n \subseteq \mathbb{R}$ be pairwise disjoint intervals.

The average value of p over I_j is

$$\overline{p_j} := \frac{1}{|I_j|} \int_{I_j} p(t)\, dt.$$

Define the linear function

$$T \colon X \longrightarrow U, \qquad Tp = (\overline{p_1}, \dots, \overline{p_n}).$$

Suppose $Tp = 0$. Then $\overline{p_j} = 0$ for all j, and so any nonzero p must change sign in I_j.

But this would imply that p has n distinct roots, which is impossible.

Thus, $N_T = \{0\}$, and so $R_T = U$.

Systems of equations

Our next two applications will rely on the following result from the previous lecture.

Example B

Take $X = U = \mathbb{R}^n$, and $T\colon \mathbb{R}^n \to \mathbb{R}^n$ given by $\displaystyle\sum_{j=1}^{n} t_{ij}x_j = u_i$, for $i = 1, \dots, n$.

If the related homogeneous system of equations $\displaystyle\sum_{j=1}^{n} t_{ij}x_j = 0$, for $i = 1, \dots, n$, has only the trivial solution $x_1 = \cdots x_n = 0$, then the inhomogeneous system T has a unique solution.

Recall that this followed from:

Corollary B

Suppose $\dim X = \dim U$ and the only vector satisfying $Tx = 0$ is $x = 0$. Then $R_T = U$.

ODEs: the method of undetermined coefficients

Consider the differential equation

$$\underbrace{ay'' + by' + cy}_{\text{homogeneous part}} \;=\; \underbrace{5e^{3t}\cos 4t}_{\text{"forcing term", } f(t)}$$

In an ODEs class, you learn that the general solution has the form $y(t) = y_h(t) + y_p(t)$.

Here, $y_h(t)$ is the general solution to the homogeneous equation $ay'' + by' + cy = 0$, i.e., the nullspace of

$$L \colon \mathcal{C}^\infty(\mathbb{R}) \longrightarrow \mathcal{C}^\infty(\mathbb{R}), \qquad L \colon y \longmapsto ay'' + by' + cy.$$

If the forcing term $f(t) = 5e^{3t}\cos 4t$ doesn't solve the homogeneous equation, we can find a "particular solution" of the form $y_p(t) = Ae^{3t}\cos 4t + Be^{3t}\sin 4t$.

But *why* does this work? Let $X = \text{Span}(e^{3t}\cos 4t,\, e^{3t}\sin 4t)$.

The only solution to the homogeneous equation $Ly = 0$ in X is $y = 0$.

We are trying to solve the inhomogeneous equation $Ly = f$, and $f \in X$.

By Example B, there is a unique $y_p \in X$ satisfying $Ly_p = f$.

PDEs: numerical solutions to Laplace's equation

Laplace's equation is $\Delta u = u_{xx} + u_{yy} = 0$, where $\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}$ is a linear operator.

Solutions to Laplace's PDE ("harmonic functions") are the functions in the nullspace of Δ.

If we fix the value of u on the boundary of a region $G \subset \mathbb{R}^2$, the solution to the boundary value problem $\Delta u = 0$ is as "flat as possible". [*Think*: plastic wrap stretched around ∂G.]

This models steady-state solutions to the heat equation PDE: $u_t = \Delta u$.

The finite difference method is a way to solve $\Delta u = 0$ numerically, using a square lattice with mesh spacing $h > 0$.

At a fixed lattice point O, let u_0 be the value of u at O, and u_W, u_E, u_N, u_S be the values at the neighbors.

We can approximate the derivatives with *centered differences*:

$$u_{xx} \approx \frac{u_W - 2u_0 + u_E}{h^2}, \qquad u_{yy} \approx \frac{u_N - 2u_0 + u_S}{h^2}.$$

Plugging this back into $\Delta u = 0$ gives $u_0 = \dfrac{u_W + u_N + u_E + u_S}{4}$, i.e., u_0 is the average of its four neighbors.

Numerical solutions to Laplace's equation (contin.)

Recall that we are trying to solve an inhomogeneous boundary value problem for Laplace's equation

$$\Delta u = 0\,, \qquad u|_{\partial G} = f(x,y) \neq 0.$$

Claim

The homogeneous equation: $\Delta u = 0$, where $u = 0$ on ∂G, has *only* the trivial solution $u_0 = 0$ for all $(x,y) \in G$.

Proof (sketch)

Let $\hat{O}$ be the lattice point at which u achieves its maximum value.

Since $u_0 = \dfrac{u_W + u_N + u_E + u_S}{4}$, then $u_0 = u_W = u_N = u_E = u_S$.

Repeating this, we see that *all* lattice points take the same value for u, and so $u = 0$.

By the result in Example B, the related inhomogeneous system for $\Delta u = 0$, with arbitrary (non-zero) boundary conditions has a unique solution. □

The algebra of linear maps

Definition

Let $S, T\colon X \to U$ be linear maps. Define

- $T + S$ by $(T + S)(x) = Tx + Sx$ for each $x \in X$.
- aT by $(aT)(x) = T(ax)$ for each $x \in X$, $a \in K$.

Easy fact

The set of linear maps from $X \to U$, denoted $\mathrm{Hom}(X, U)$, or $\mathscr{L}(X, U)$, is a vector space.

Lemma 2.3 (HW)

If $T\colon X \to U$ and $S\colon U \to V$ are linear maps, then so is $(S \circ T)\colon X \to V$.

Moreover, composition is distributive w.r.t. addition. That is, if $P, T\colon X \to U$ and $R, S\colon U \to V$, then

$$(R + S) \circ T = R \circ T + S \circ T, \qquad S \circ (T + P) = S \circ T + S \circ P\,.$$

Remarks

- We usually just write $S \circ T$ as just ST.
- In general, $ST \neq TS$ (note that TS may not even be defined).

Invertibility

Definition

A linear map T is invertible if it is 1–1 and onto (i.e., if it is an isomorphism). Denote the inverse by T^{-1}.

Exercise

If T is invertible, then TT^{-1} is the identity.

Proposition 2.4 (exercise)

Let $T\colon X \to U$ be linear.

(i) If T is linear, then so is T^{-1}.

(ii) If S and T are invertible and ST defined, then it is invertible with $(ST)^{-1} = T^{-1}S^{-1}$.

Examples

(ix) Take $X = U = V = \mathbb{R}[t]$, with $T = \frac{d}{dt}$ and $S =$ multiplication by t.

(x) Take $X = U = V = \mathbb{R}^3$, with S a 90°-rotation around the x_1 axis, and T a 90°-rotation around the x_2 axis.

In both of these examples, S and T are linear with $ST \neq TS$. (Which are invertible?)

More on the algebra of linear maps

Definition

An endomorphism of X is a linear map from X to itself. Denote the set of endomorphisms of X by $\mathrm{Hom}(X,X)$ or $\mathscr{L}(X,X)$ or $\mathrm{End}(X)$.

Remarks

$\mathrm{Hom}(X,X)$ is a vector space, but we can also "multiply" vectors; it is an algebra.

It is an associative but noncommutative algebra, with unity I, satisfying $Ix = x$.

$\mathrm{Hom}(X,X)$ contains zero divisors: pairs S, T such that $ST = 0$ but neither S nor T is zero.

Proposition

If $A \in \mathrm{Hom}(X,X)$ is a left inverse of $B \in \mathrm{Hom}(X,X)$ [i.e., $AB = I$], then it is also a right inverse [i.e., $BA = I$]. □

Definition

The invertible elements of $\mathrm{Hom}(X,X)$ forms the general linear group, denoted $\mathrm{GL}_n(K)$, where $n = \dim X$.

Every $S \in \mathrm{GL}_n(K)$ defines a similarity transformation ϕ_S of $\mathrm{Hom}(X,X)$, sending $M \longmapsto M_S := SMS^{-1}$, for each $M \in \mathrm{Hom}(X,X)$. We say M and M_S are similar.

Similarity

Proposition 2.5

Every similarity transform is an automorphism ["structure-preserving bijection"] of $\mathrm{Hom}(X, X)$:

$$(kM)_S = kM_S, \qquad (M+N)_S = M_S + N_S, \qquad (MN)_S = M_S N_S\,.$$

Moreover, the set of similarity transforms forms a group under $(M_S)_T := M_{TS}$, called the inner automorphism group of $\mathrm{GL}_n(K)$.

Proof

Proposition 2.6 (exercise)

Similarity is an equivalence relation, i.e., it is:

(i) Reflexive: $M \sim M$;

(ii) Symmetric: $L \sim M$ implies $M \sim L$;

(iii) Transitive: $L \sim M$ and $M \sim N$ implies $L \sim N$. □

More on the algebra of linear maps

Proposition 2.7

If either A or B in $\mathrm{Hom}(X,X)$ is invertible, then AB and BA are similar. □

Given any $A \in \mathrm{Hom}(X,X)$ and polynomial $p(t) = a_N t^N + \cdots + a_1 t + a_0$, consider the polynomial $p(A) = a_N A^N + \cdots + a_1 A + a_0 I$.

The set of polynomials in A is a commutative subalgebra of $\mathrm{Hom}(X,X)$. [to be revisited]

Miscellaneous definitions

- A linear map $P \colon X \to X$ is a projection if $P^2 = P$.
- The commutator of $A, B \in \mathrm{Hom}(X,X)$ is $[A,B] := AB - BA$, which is 0 iff A and B commute.

Examples (contin.)

(xii) If $X = \{f : \mathbb{R} \to \mathbb{R},\ \text{contin.}\}$, then the following maps $P, Q \in \mathrm{Hom}(X,X)$ are projections:

- $(Pf)(x) = \dfrac{f(x) + f(-x)}{2}$; this is the even part of f.
- $(Qf)(x) = \dfrac{f(x) - f(-x)}{2}$; this is the odd part of f.

Note that $f = Pf + Qf$ for any $f \in X$.

Overview

Soon, we will learn about the transpose of a linear map, and then how to encode an arbitrary linear map with a matrix.

Though it is not necessary, it is helpful to have some familiarity with undergraduate-level matrix analysis *before* seeing these.

Let's review now the "four subspaces" that arise from every matrix:

1. column space
2. row space
3. nullspace
4. left nullspace

Understanding these subspaces will motivate the more theoretical concepts and results as they arise, and give them context.

Throughout, we strongly recommend viewing:

1. vectors $x \in X$ as column vectors
2. scalar functions $\ell \in X'$ as row vectors.

Column space, and row space

Let $A\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear map, which we can think of as an $m \times n$ matrix, $A = (a_{ij})$.

The transpose is a linear map $A^T : \mathbb{R}^m \to \mathbb{R}^n$, which we can think of as an $n \times m$ matrix.

Definition

The range R_A is the span of the column vectors, called the column space of A.

Its dimension is called the column rank of A.

The range R_{A^T} is the span of the column vectors of A^T, called the row space of A.

Its dimension is called the row rank of A.

Theorem

$\dim R_A = \dim R_{A^T}$, which we call the rank of A.

Moreover, the restriction of $A\colon R_{A^T} \to R_A$ is bijective.

Column space, row space, nullspace and left nullspace

(picture on board)

Four ways to multiply matrices

Suppose we have linear maps $\mathbb{R}^p \xrightarrow{B} \mathbb{R}^n \xrightarrow{A} \mathbb{R}^m$. As matrices, we can multiply them:

1. rows by columns
2. by columns
3. by rows
4. columns by rows

Systems of equations and Gaussian elimination

Let's review how to solve a system of equations, and how it relates to the 4 subspaces.

$$
\begin{aligned}
x_1 + x_2 + 2x_3 + 3x_4 &= u_1 \\
x_1 + 2x_2 + 3x_3 + x_4 &= u_2 \\
2x_1 + x_2 + 2x_3 + 3x_4 &= u_3 \\
3x_1 + 4x_2 + 6x_3 + 2x_4 &= u_4
\end{aligned}
$$

The transpose of a linear map

Every undergraduate linear algebra student learns about the transpose of a matrix, formed by flipping it across its main diagonal.

But what does this *really* represent?

The transpose of a matrix is what results from swapping rows with columns.

In our setting, we like to think about vectors in X as column vectors, and dual vectors in X' as row vectors.

The transpose is a more general concept than just an operation on matrices.

Given a linear map $T\colon X \to U$, its transpose is a certain induced linear map $T'\colon U' \to X'$ between the dual spaces.

Now we'll learn how to encode linear maps with matrices. When we do this, the matrix of the transpose map will simply be the tranpose of the matrix.

Let's start with the definition of transpose of a linear map and then learn about some basic properties.

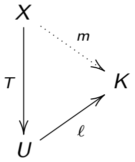
The transpose of a linear map

Let $T\colon X \to U$ be linear and $\ell \in U'$.

The composition $m := \ell T$ is a linear map $X \to K$.

Since T is fixed, this defines a linear map, called the transpose of T:

$$T'\colon U' \longrightarrow X', \qquad T'\colon \ell \longmapsto m,$$



Using scalar product notation we can rewrite $m(x) = \ell(T(x))$ as $(m,x) = (\ell, Tx)$.

Key property

The transpose of $T\colon X \to U$ is the (unique) map $T'\colon U' \to X'$ that satisfies $m = T'\ell$, i.e.,

$$(T'\ell, x) = (\ell, Tx), \qquad \text{for all } x \in X,\ \ell \in U'.$$

Caveat: We are writing ℓT for $\ell \circ T$, but $T'\ell$ for $T'(\ell)$ (much like Tx for $T(x)$).

Properties (HW exercise)

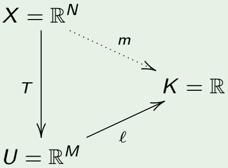
Whenever meaningful, we have

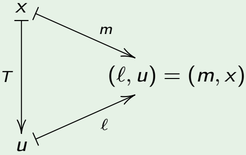
$$(ST)' = T'S', \qquad (T+R)' = T' + R', \qquad (T^{-1})' = (T')^{-1}.$$

Another example of a linear map, and its transpose

Examples (cont.)

(xi) Let $X = \mathbb{R}^N$, $U = \mathbb{R}^M$, and $Tx = u$, where $u_i = \sum_{j=1}^{N} t_{ij} x_j$.





By definition, for some $\ell_1, \dots, \ell_m \in K$,

$$(\ell, u) = \sum_{i=1}^{M} \ell_i u_i = \sum_{i=1}^{M} \ell_i \left(\sum_{j=1}^{N} t_{ij} x_j \right) = \sum_{i=1}^{M} \sum_{j=1}^{N} \ell_i t_{ij} x_j = \sum_{i=1}^{N} \left(\ell_i \sum_{j=1}^{M} t_{ij} x_j \right) = \sum_{j=1}^{N} m_j x_j$$

This gives us a formula for $m = (m_1, \dots, m_N)$, where $(\ell, u) = (m, x)$.

We'll see later that if we express T in matrix form, then T' is formed by making the rows of T the columns of T'.

What does this really mean?

$$(\ell, u) = \sum_{i=1}^{M} \ell_i u_i = \sum_{i=1}^{M} \ell_i \left(\sum_{j=1}^{N} t_{ij} x_j \right) = \sum_{i=1}^{M} \sum_{j=1}^{N} \ell_i t_{ij} x_j = \sum_{i=1}^{N} \left(\ell_i \sum_{j=1}^{M} t_{ij} x_j \right) = \sum_{j=1}^{N} m_j x_j$$

The nullspace of the transpose

Proposition 2.8

If X'' and U'' are canonically identified with X and U, respectively, then $T'' = T$. $\square$

Proposition 2.9

The annihilator of the range of T is the nullspace of its transpose, i.e., $R_T^\perp = N_{T'}$.

Proof

Applying $\perp$ to both sides of $R_T^\perp = N_{T'}$ (Proposition 2.9) yields the following:

Corollary 2.10

The range of T is the annihilator of the nullspace of T', i.e., $R_T = N_{T'}^\perp$. $\square$

The rank of the transpose

Theorem 2.11

For any linear mapping $T\colon X \to U$, we have $\dim R_T = \dim R_{T'}$.

Proof

Corollary 2.12

Let $T\colon X \to U$ be linear with $\dim X = \dim U$. Then $\dim N_T = \dim N_{T'}$.

Proof

How to encode a linear map with a matrix

Let $T \colon X \to U$ be a linear map between finite-dimensional vector spaces.

To encode T as a matrix, we'll need to choose:

1. an "input basis" $\mathcal{B}_X = \{x_1, \dots, x_n\}$ for X,
2. an "output basis" $\mathcal{B}_U = \{u_1, \dots, u_m\}$ for U.

Let $\{\ell_1, \dots, \ell_m\}$ be the dual basis of $\mathcal{B}_U$.

First, we write the images of the basis vectors in $\mathcal{B}_X$ using the basis vectors in $\mathcal{B}_U$:

$$
\begin{aligned}
Tx_1 &= \\
Tx_2 &= \\
&\vdots \\
Tx_j &= \\
&\vdots \\
Tx_n &=
\end{aligned}
$$

Summary of how to write a linear maps as a matrix

Let $T\colon X \to U$. The matrix A of T w.r.t. bases $\mathcal{B}_X = \{x_1, \dots, x_n\}$ and $\mathcal{B}_U = \{u_1, \dots, u_m\}$ is

$$A = {}_{\mathcal{B}_X}\big[T\big]_{\mathcal{B}_U} = \Big[Tx_1 \;\; Tx_2 \;\; \cdots \;\; Tx_n\Big].$$

Remarks

- The range of T is the span of the column vectors – the column space.
- $a_{ij} = (\ell_i, Tx_j)$,

$$\begin{aligned}
Tx_1 &= a_{11}u_1 + a_{21}u_1 + \cdots + a_{i1}u_j + \cdots + a_{m1}u_m \\
Tx_2 &= a_{12}u_1 + a_{22}u_1 + \cdots + a_{i2}u_j + \cdots + a_{m2}u_m \\
&\vdots \\
Tx_j &= a_{1j}u_1 + a_{2j}u_1 + \cdots + a_{ij}u_j + \cdots + a_{mj}u_m \\
&\vdots \\
Tx_n &= a_{1n}u_1 + a_{2n}u_1 + \cdots + a_{in}u_j + \cdots + a_{mn}u_m
\end{aligned}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Example 1

Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the projection onto the line $y = x$.

An interesting choice of basis

Proposition

If $T \colon X \to U$ is invertible, we can always choose $\mathcal{B}_X$ and $\mathcal{B}_U$ so the matrix is the identity.

More generally, for any $T \colon X \to U$, we can choose $\mathcal{B}_X$ and $\mathcal{B}_U$ so the matrix in block form is

$$A = {}_{\mathcal{B}_X}\big[T\big]_{\mathcal{B}_U} = \begin{bmatrix} I_{r \times r} & 0 \\ 0 & 0 \end{bmatrix}.$$

Example 2

Let $X = \{c_0 + c_1 x + c_2 x^2 \mid c_i \in \mathbb{R}\}$ with basis $\mathcal{B}_X = \{1, x, x^2\}$.

Let $U = \{c_0 + c_1 x \mid c_i \in \mathbb{R}\}$ with basis $\mathcal{B}_U = \{1, x\}$.

Let $T = \dfrac{d}{dx}$, and so $T\colon c_0 + c_1 x + c_2 x^2 \mapsto c_1 + 2c_2 x$.

The matrix of the transpose

Let $T\colon X \to U$ be linear, and pick bases $\mathcal{B}_X = \{x_1, \dots, x_n\}$ and $\mathcal{B}_U = \{u_1, \dots, u_m\}$.

Let $\mathcal{B}_{U'} = \{\ell_1, \dots, \ell_m\}$ be the dual basis of $\mathcal{B}_U$.

Let $A = (a_{ij})$ be the matrix of T w.r.t. these bases.

In plain English, a_{ij} is the result of:

1. starting with the j^{th} basis vector in X,
2. applying the map T,
3. applying the i^{th} dual basis vector in U'.

Let's apply these steps to the transpose map $T'\colon U' \to X'$ to find its matrix form, $A' = (a'_{ij})$.

How to encode a linear map with a matrix

Let $T\colon X \to U$ be a linear map between finite-dimensional vector spaces.

To encode T as a matrix, we'll need to choose:

1. an "input basis" $\mathcal{B}_X = \{x_1, \dots, x_n\}$ for X,
2. an "output basis" $\mathcal{B}_U = \{u_1, \dots, u_m\}$ for U.

Let $\{\ell_1, \dots, \ell_m\}$ be the dual basis of $\mathcal{B}_U$.

First, we write the images of the basis vectors in $\mathcal{B}_X$ using the basis vectors in $\mathcal{B}_U$:

$$\begin{aligned} Tx_1 &= \\ Tx_2 &= \\ &\vdots \\ Tx_j &= \\ &\vdots \\ Tx_n &= \end{aligned}$$

Summary

Let $T: X \to U$. The matrix A of T w.r.t. bases $\mathcal{B}_X = \{x_1, \dots, x_n\}$ and $\mathcal{B}_U = \{u_1, \dots, u_m\}$ is

$$A = {}_{\mathcal{B}_X}\big[T\big]_{\mathcal{B}_U} = \Big[Tx_1 \quad Tx_2 \quad \cdots \quad Tx_n\Big].$$

Remarks

- The range of T is the span of the column vectors – the column space.
- $a_{ij} = (\ell_i, Tx_j)$,

$$\begin{aligned}
Tx_1 &= a_{11}u_1 + a_{21}u_1 + \cdots + a_{i1}u_j + \cdots + a_{m1}u_m \\
Tx_2 &= a_{12}u_1 + a_{22}u_1 + \cdots + a_{i2}u_j + \cdots + a_{m2}u_m \\
&\vdots \\
Tx_j &= a_{1j}u_1 + a_{2j}u_1 + \cdots + a_{ij}u_j + \cdots + a_{mj}u_m \\
&\vdots \\
Tx_n &= a_{1n}u_1 + a_{2n}u_1 + \cdots + a_{in}u_j + \cdots + a_{mn}u_m
\end{aligned}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Example 1

Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be the projection onto the line $y = x$.

An interesting choice of basis

Proposition

If $T\colon X \to U$ is invertible, we can always choose $\mathcal{B}_X$ and $\mathcal{B}_U$ so the matrix is the identity.

More generally, for any $T\colon X \to U$, we can choose $\mathcal{B}_X$ and $\mathcal{B}_U$ so the matrix in block form is

$$A = {}_{\mathcal{B}_X}\big[T\big]_{\mathcal{B}_U} = \begin{bmatrix} I_{r\times r} & 0 \\ 0 & 0 \end{bmatrix}.$$

Example 2

Let $X = \{c_0 + c_1 x + c_2 x^2 \mid c_i \in \mathbb{R}\}$ with basis $\mathcal{B}_X = \{1, x, x^2\}$.

Let $U = \{c_0 + c_1 x \mid c_i \in \mathbb{R}\}$ with basis $\mathcal{B}_U = \{1, x\}$.

Let $T = \dfrac{d}{dx}$, and so $T \colon c_0 + c_1 x + c_2 x^2 \mapsto c_1 + 2c_2 x$.

The matrix of the transpose

Let $T\colon X \to U$ be linear, and pick bases $\mathcal{B}_X = \{x_1, \ldots, x_n\}$ and $\mathcal{B}_U = \{u_1, \ldots, u_m\}$.

Let $\mathcal{B}_{U'} = \{\ell_1, \ldots, \ell_m\}$ be the dual basis of $\mathcal{B}_U$.

Let $A = (a_{ij})$ be the matrix of T w.r.t. these bases.

In plain English, a_{ij} is the result of:

1. starting with the j^{th} basis vector in X,
2. applying the map T,
3. applying the i^{th} dual basis vector in U'.

Let's apply these steps to the transpose map $T'\colon U' \to X'$ to find its matrix form, $A' = (a'_{ij})$.

Change of basis

Previously, we learned how a linear map $T\colon X \to U$ is encoded by a matrix, with respect to an input basis $\mathcal{B}_X$ and output basis $\mathcal{B}_U$.

It is natural to ask how changing the bases changes the matrix.

We will answer this question now.

In the special case of $T\colon X \to X$, we will see that two matrices A and B can represent the same linear map if they are similar. That is,

$$A = PBP^{-1}, \qquad \text{for some invertible matrix } P.$$

We will show to how construct such a P, which is called a change of basis matrix.

Change of basis matrices

Let $T\colon X \to U$ be linear, and $x_1, \dots, x_n$ and $u_1, \dots, u_m$ be bases.

Since $\dim X = n$ and $\dim U = m$, we have $X \cong K^n$ and $U \cong K^m$. (Let's say $K = \mathbb{R}$.)

An example in $\mathbb{R}^2$

Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be linear, and A the 2×2 matrix w.r.t. the standard basis $e_1, e_2 \in \mathbb{R}^2$.

Let's see what the matrix is with respect to a different basis, $v_1 = \begin{bmatrix} a \\ c \end{bmatrix}$ and $v_2 = \begin{bmatrix} b \\ d \end{bmatrix}$.