Mathematics for Computer Science
Linear Algebra

Lecture 20: General linear transformations

Andrei Krokhin

April 18, 2021

Linear maps

Definition

Let V and W be vector spaces. A function $T : V \to W$ is called a linear map, or a linear transformation from V to W if, for all $\mathbf{u}, \mathbf{v} \in V, k \in \mathbb{R}$,

$$T(\mathbf{u} + \mathbf{v}) = T(\mathbf{u}) + T(\mathbf{v}) \text{ and } T(k\mathbf{u}) = k \cdot T(\mathbf{u}).$$

If $V = W$ then T is called a linear operator on V.

Recall:

- The linear maps from $\mathbb{R}^n$ to $\mathbb{R}^m$ are precisely the matrix transformations, i.e. the maps of the form $T_A(\mathbf{x}) = A\mathbf{x}$, where A is an $m \times n$ matrix.
- This can be generalised to all finite-dimensional vector spaces V and W.
- Key observation: if $B = \{\mathbf{v}_1, \ldots, \mathbf{v}_n\}$ is a basis in V and T is a linear map from V, then $\mathbf{v} = k_1\mathbf{v}_1 + \ldots + k_n\mathbf{v}_n$ implies $T(\mathbf{v}) = k_1 T(\mathbf{v}_1) + \ldots + k_n T(\mathbf{v}_n)$.

  To rephrase: images of basis vectors uniquely determine a linear map.

Examples: inner product and evaluation maps

Let V be an inner product space. Fix any vector $\mathbf{v}_0 \in V$ and consider the map

$$T(\mathbf{x}) = \langle \mathbf{x}, \mathbf{v}_0 \rangle.$$

This is a linear map from V to $\mathbb{R}^1$.

Indeed, the properties of inner product guarantee that

$$T(k\mathbf{u}) = \langle k\mathbf{u}, \mathbf{v}_0 \rangle = k\langle \mathbf{u}, \mathbf{v}_0 \rangle = kT(\mathbf{u}),$$

$$T(\mathbf{u} + \mathbf{v}) = \langle \mathbf{u} + \mathbf{v}, \mathbf{v}_0 \rangle = \langle \mathbf{u}, \mathbf{v}_0 \rangle + \langle \mathbf{v}, \mathbf{v}_0 \rangle = T(\mathbf{u}) + T(\mathbf{v}).$$

Consider the space $F(-\infty, \infty)$ of all functions from $\mathbb{R}$ to $\mathbb{R}$. Fix sample points $x_1, \ldots, x_n$ and consider the (evaluation) map

$$T(f) = (f(x_1), \ldots, f(x_n)).$$

It is straightforward to check that this is a linear map from $F(-\infty, \infty)$ to $\mathbb{R}^n$.

Examples: differentiation and integration maps

Let $C^1(-\infty, \infty)$ be the space of all functions $f \in F(-\infty, \infty)$ with continuous first derivative f'. Then the following is a linear map from $C^1(-\infty, \infty)$ to $C(-\infty, \infty)$:

$$D(f) = f'.$$

Indeed, by rules of differentiation, we have

$$D(f+g) = (f+g)' = f'+g' = D(f)+D(g) \quad \text{and} \quad D(kf) = (kf)' = kf' = kD(f).$$

The following map from $C(-\infty, \infty)$ to $C^1(-\infty, \infty)$ is also linear:

$$J(f) = \int_0^x f(t)\, dt.$$

For example, if $f(x) = x^2$ then $J(f) = \int_0^x t^2\, dt = \frac{t^3}{3}\Big|_0^x = \frac{x^3}{3}$.

The linearity of the map J follows immediately from the rules of integration.

Kernel and range of a linear map

Definition

If $T : V \to W$ is a linear map then

- the set $ker(T) = \{\mathbf{x} \in V \mid T(\mathbf{x}) = \mathbf{0}\}$ is called the kernel of T, and
- the set $range(T) = \{\mathbf{y} \in W \mid T(\mathbf{x}) = \mathbf{y} \text{ for some } \mathbf{x} \in V\}$ is the range of T.

Exercise: Prove that $ker(T)$ is a subspace of V and $range(T)$ is a subspace of W.

Example: For a matrix transformation $T_A : \mathbb{R}^n \to \mathbb{R}^m$,

- $ker(T_A)$ is the null space of A, and $dim(ker(T_A)) = nullity(A)$, and
- $range(T_A)$ is the column space of A, and $dim(range(T_A)) = rank(A)$.

Recall the dimension theorem for matrices:

- if a matrix A has size $m \times n$ then $rank(A) + nullity(A) = n$.

Theorem (Dimension theorem for linear maps)

For any linear map $T : V \to W$, if V is finite-dimensional then

$$dim(ker(T)) + dim(range(T)) = dim(V).$$

Example

One corollary from the dimension theorem is this:

For any linear operator T on $\mathbb{R}^n$, TFAE:

1. T is injective (equivalently, $ker(T) = \{\mathbf{0}\}$),
2. T is surjective, i.e. $range(T) = \mathbb{R}^n$.

Both conditions are equivalent to the invertibility of A, where A is s.t. $T = T_A$.

Consider the vector space $\mathbb{R}^\infty$ of all infinite sequences $(x_1, x_2, \ldots, x_n, \ldots)$.

Do the implications (1) $\Rightarrow$ (2) and (2) $\Rightarrow$ (1) hold for linear operators on $\mathbb{R}^\infty$?

No, consider the "shift" operators T_ℓ and T_r:

$$\begin{aligned} T_\ell(x_1, x_2, \ldots, x_n, \ldots) &= \quad (x_2, x_3, \ldots, x_n, \ldots) \\ T_r(x_1, x_2, \ldots, x_n, \ldots) &= (0, x_1, x_2, \ldots, x_n, \ldots) \end{aligned}$$

Example

Consider the differentiation map $D : C^1(-\infty, \infty) \to C(-\infty, \infty)$.

- What is $ker(D)$?

  We have $D(f) = f' = 0$ iff f is a constant function.
  Hence, $ker(D)$ is the set of all constant functions.

  One basis of $ker(D)$ is $\{1\}$, the constant function $f(x) = 1$.

- What is $range(D)$?

  We know from calculus that, for any function $f \in C(-\infty, \infty)$, we have

$$\frac{d}{dx} \int_0^x f(t)\, dt = f(x).$$

  Hence, each function $f \in C(-\infty, \infty)$ can be represented as

$$f = D(g) \text{ where } g(x) = \int_0^x f(t)\, dt, \text{ i.e. } g = J(f).$$

  So, $range(D) = C(-\infty, \infty)$.

Eigenvalues and eigenvectors

Definition

Let T be a linear operator on a vector space V. A non-zero vector $\mathbf{v} \in V$ is called an eigenvector of T if $T(\mathbf{v}) = \lambda \mathbf{v}$ for some scalar λ. In this case, λ is called an eigenvalue of T and $\mathbf{v}$ is an eigenvector corresponding to λ.

Example: Let $C^\infty(-\infty, \infty)$ be the space of all infinitely differentiable functions (i.e. functions f such that the n-th derivative $f^{(n)}$ exists for all n) and consider the differentiation operator D on $C^\infty(-\infty, \infty)$.

- A number λ is an eigenvalue of D if $D(f) = f' = \lambda f$ for some function $f \neq 0$.
- Solving the differential equation $f'(x) = \lambda f(x)$, get $f(x) = ce^{\lambda x}$, c constant
- Hence, every number $\lambda \in \mathbb{R}$ is an eigenvalue of D, with a corresponding eigenvector $f(x) = e^{\lambda x}$.
- Moreover, each eigenspace of D is 1-dimensional, with basis $e^{\lambda x}$.

Matrices for linear transformations

We know:

- If T is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $\mathbf{e}_1, \ldots, \mathbf{e}_n$ is the standard basis in $\mathbb{R}^n$, then $T = T_A$, i.e. $T(\mathbf{x}) = A\mathbf{x}$, where

$$A = [T(\mathbf{e}_1)| \ldots | T(\mathbf{e}_n)].$$

How do we generalise this from standard to arbitrary bases?

- Let $T : V \to W$ be a linear map, and let $dim(V) = n$ and $dim(W) = m$.
- Fix a basis $B = \{\mathbf{u}_1, \ldots, \mathbf{u}_n\}$ in V and a basis $B' = \{\mathbf{v}_1, \ldots, \mathbf{v}_m\}$ in W.
- Let

$$[T]_{B',B} = [\,(T(\mathbf{u}_1))_{B'} \mid \ldots \mid (T(\mathbf{u}_n))_{B'}\,].$$

- This matrix $[T]_{B',B}$ is called the matrix of T relative to the bases B and B'.
- It can be checked directly that $[T]_{B',B}(\mathbf{x})_B = (T(\mathbf{x}))_{B'}$ for any $\mathbf{x} \in V$, where $(\mathbf{x})_B$ is the coordinate vector of $\mathbf{x}$ in basis B, and similarly for $(T(\mathbf{x}))_{B'}$.

Example

Consider the linear map $T : \mathbb{R}^2 \to \mathbb{R}^3$ given by

$$T(x_1, x_2) = (x_2,\ -5x_1 + 13x_2,\ -7x_1 + 16x_2).$$

What is the matrix of T relative to the bases $B = \{\mathbf{u}_1 = (3,1), \mathbf{u}_2 = (5,2)\}$ and $B' = \{\mathbf{v}_1 = (1,0,-1), \mathbf{v}_2 = (-1,2,2), \mathbf{v}_3 = (0,1,2)\}$?

A direct computation shows that $T(\mathbf{u}_1) = (1,-2,-5)$ and $T(\mathbf{u}_2) = (2,1,-3)$.
Expressing these vectors via B', we get (check!)

$$T(\mathbf{u}_1) = \mathbf{v}_1 - 2\mathbf{v}_3 \text{ and } T(\mathbf{u}_2) = 3\mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3.$$

Hence, we have

$$[T]_{B',B} = [\,(T(\mathbf{u}_1))_{B'} \mid (T(\mathbf{u}_2))_{B'}\,] = \begin{pmatrix} 1 & 3 \\ 0 & 1 \\ -2 & -1 \end{pmatrix}.$$

Change of basis

Consider the special case when $V = W$ and let B, B' be two bases of V.

Apply the formula

$$[T]_{B',B} = [\,(T(\mathbf{u}_1))_{B'} \mid \dots \mid (T(\mathbf{u}_n))_{B'}\,].$$

to the linear map $T = I$, i.e. the identity map on V ($I(\mathbf{x}) = \mathbf{x}$). Then we have

$$[I]_{B',B} = [\,(\mathbf{u}_1)_{B'} \mid \dots \mid (\mathbf{u}_n)_{B'}\,].$$

This matrix, usually denoted by $P_{B \to B'}$, corresponds to the change of basis in V.

It is called a transition matrix from B to B' and we have $P_{B \to B'}(\mathbf{x})_B = (\mathbf{x})_{B'}$.

An algorithm for computing $P_{B \to B'}$:

- Form the matrix $[B'|B]$.
- Use elementary row operations to transfrom it to reduced row echelon form.
- The resulting matrix is $[I|P_{B \to B'}]$

Example

Find the transition matrices $P_{B\to B'}$ and $P_{B'\to B}$ for the following bases in $\mathbb{R}^2$:

$$B = \{\mathbf{u}_1 = (1,0),\ \mathbf{u}_2 = (0,1)\} \text{ and } B' = \{\mathbf{u}_1' = (1,1),\ \mathbf{u}_2' = (2,1)\}.$$

We have

$$[B'|B] = \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 2 & 1 & 0 \\ 0 & -1 & -1 & 1 \end{array}\right) \rightsquigarrow \left(\begin{array}{cc|cc} 1 & 0 & -1 & 2 \\ 0 & 1 & 1 & -1 \end{array}\right)$$

Hence,

$$P_{B\to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \text{ and we also have } P_{B'\to B} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}$$

What are the coordinates of $\mathbf{x} = (2,-3)$ relative to the basis B' ?

Since B is the standard basis, we have $(\mathbf{x})_B = \mathbf{x}$, so

$$(\mathbf{x})_{B'} = P_{B\to B'}(\mathbf{x})_B = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix}\begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} -8 \\ 5 \end{pmatrix}$$

Change of basis and similarity of matrices

For a linear operator T on V and a basis B of V, we write $[T]_B$ instead of $[T]_{B,B}$.

If B and B' are different bases of V, how are the matrices $[T]_B$ and $[T]_{B'}$ related?

Theorem

Let $T : V \to V$ be a linear operator and let B and B' be bases of V. Then

$$[T]_{B'} = P^{-1}[T]_B P$$

where $P = P_{B' \to B}$ and $P^{-1} = P_{B \to B'}$.

Recall: any matrices related as $[T]_B$ and $[T]_{B'}$ above are called *similar*.

Any similarity relationship $A' = P^{-1}AP$ can be interpreted in this way:

If $A = [T]_B$ for abasis B then $A' = [T]_{B'}$ for another basis B' and $P = P_{B' \to B}$.

From this perspective, diagonalisation of a square matrix A = search for a basis B' such that the matrix $[T_A]_{B'}$ (of the operator T_A relative to basis B') is diagonal.

Example

Consider the linear operator $T = T_A$ on $\mathbb{R}^2$ where

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 4 \end{pmatrix}.$$

In other words, $A = [T]_B$ where B is the standard basis in $\mathbb{R}^2$.

Consider the basis $B' = \{\mathbf{u}'_1 = (1,1),\ \mathbf{u}'_2 = (2,1)\}$ (same as in previous example).

We found in the previous example:

$$P_{B \to B'} = \begin{pmatrix} -1 & 2 \\ 1 & -1 \end{pmatrix} \text{ and } P_{B' \to B} = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix}.$$

From the previous slide, we have

$$[T]_{B'} = P_{B \to B'}[T]_B P_{B' \to B} = P_{B \to B'} A P_{B' \to B} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix}.$$

In particular, B' is a basis that diagonalises the matrix A.

What we learnt today

- General linear maps
- Kernel, range, and eigenvalues/vectors of a linear map
- Matrix of a linear map relative to bases
- Change of basis and similarity of matrices

Gilbert Strang's "big picture" of Linear Algebra