Linear Transformations

Bruno Carpentieri
bruno.carpentieri@unibz.it

3rd floor - Room 3.10 - Phone +39 047 101 6027

Course: Linear Algebra
Academic year: 2023-2024

Version: January 10, 2024

Notes

Outline

Linear Operators

The Matrix of a Transformation

Composite Transformations

Notes

Outline

Linear Operators

The Matrix of a Transformation

Composite Transformations

Notes

Idea of Linear Transformation

- It is natural to consider functions defined between vector spaces, e.g. rotations, projections, scaling, etc. They are called *transformations*.
- Given two vector spaces $(\mathcal{V}, +, \cdot)$ and $(\mathcal{W}, +, \cdot)$, a transformation T from $\mathcal{V}$ and $\mathcal{W}$ assigns an output $T(\vec{v}) \in \mathcal{W}$ to an input vector $\vec{v} \in \mathcal{V}$. The transformation is linear if it meets these requirements:
  1. $T(\vec{v} + \vec{w}) = T(\vec{v}) + T(\vec{w})$ for all $\vec{v}, \vec{w}$;
  2. $T(c\vec{v}) = cT(\vec{v})$ for all scalars c and vectors $\vec{v}$.
- We can combine 1. and 2. as:

$$\boxed{\text{Linear transformation: } T(c\vec{v} + d\vec{w}) = cT(\vec{v}) + dT(\vec{w})}$$

Notes

Linear transformations

✔ Property: Linearity extends to combinations of n vectors. When T is linear, $\vec{u} = c_1\vec{v}_1 + c_2\vec{v}_2 + \cdots + c_n\vec{v}_n \Rightarrow T(\vec{u}) = c_1 T(\vec{v}_1) + c_2 T(\vec{v}_2) + \cdots + c_n T(\vec{v}_n)$.

✔ Property: When T is linear, $T(\vec{0}) = \vec{0}$ because $T(\vec{v}) = T(\vec{v} + \vec{0}) = T(\vec{v}) + T(\vec{0})$, hence $T(\vec{0}) = T(\vec{v}) - T(\vec{v}) = \vec{0}$ for every vector $\vec{v}$.

✔ Example: The function $0(\vec{x}) = \vec{0}$ that maps all vectors in a space $\mathcal{U}$ to the zero vector in another space $\mathcal{V}$ is a linear transformation from $\mathcal{U}$ into $\mathcal{V}$, and, not surprisingly, it is called the zero transformation.

✔ Example: Matrix multiplication is linear. This $T(\vec{v}) = A\vec{v}$ is linear because

$$A(c\vec{v} + d\vec{w}) = cA\vec{v} + dA\vec{w}.$$

✔ Example: When $rank(A) = n$, another linear transformation is multiplication by A^{-1} (the *inverse transformation* T^{-1}), which brings every vector $T(\vec{v})$ back to $\vec{v}$:

$$T^{-1}(T(\vec{v})) = \vec{v} \text{ matches the matrix multiplication } A^{-1}(A(\vec{v})) = \vec{v}.$$

Notes

Further Examples

1. A linear transformation is highly restricted. Suppose T adds $\vec{u}_0$ to every vector. Then $T(\vec{v}) = \vec{v} + \vec{u}_0$ and $T(\vec{w}) = \vec{w} + \vec{u}_0$. This *isn't linear*. Applying T to $\vec{v} + \vec{w}$ produces $\vec{v} + \vec{w} + \vec{u}_0$. That is not the same as $T(\vec{v}) + T(\vec{w}) = \vec{v} + \vec{w} + 2\vec{u}_0$.

   Shift is not linear: $\vec{v} + \vec{w} + \vec{u}_0 \neq T(\vec{v}) + T(\vec{w}) = \vec{v} + \vec{u}_0 + \vec{w} + \vec{u}_0$.

2. The exception is when $\vec{u}_0 = 0$. The transformation reduces to $T(\vec{v}) = \vec{v}$. The transformation $T(\vec{v}) = \vec{v}$ is called the identity transformation (nothing moves, as in multiplication by the identity matrix). *That is certainly linear.* In this case the input space $\mathcal{V}$ is the same as the output space $\mathcal{W}$.

3. The linear-plus-shift transformation $T(\vec{v}) = A\vec{v} + \vec{u}_0$ is called affine. Straight lines stay straight although T is *not linear.* Computer graphics works with affine transformations, because we must be able to move images.

Notes

Further Examples

4. Transformations can be also defined on functional spaces. If $\mathcal{W}$ is the vector space of all functions from $\mathbb{R}$ to $\mathbb{R}$, and if $\mathcal{V}$ is the vector space of all differentiable functions from $\mathbb{R}$ to $\mathbb{R}$, then the mapping $D(f) = df/dx$ is a linear transformation from $\mathcal{V}$ into $\mathcal{W}$ because

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

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

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

Notes

Exercise

✔ Problem. Choose a fixed vector $\vec{a} = (1, 3, 4)$, and let $T(\vec{v})$ be the dot product $\vec{a} \cdot \vec{v}$:

The input is $\vec{v} = (v_1, v_2, v_3)$. The output is $T(\vec{v}) = \vec{a} \cdot \vec{v} = v_1 + 3v_2 + 4v_3$.

Determine if such T is linear.

Solution.

This T is linear. The inputs $\vec{v}$ come from three-dimensional space, so $\mathcal{V} = \mathbb{R}^3$. The outputs are just numbers, so the output space is $\mathcal{W} = \mathbb{R}^1$. We are multiplying by the row matrix $A = [\,1\ 3\ 4\,]$. Then $T(\vec{v}) = A\vec{v}$.

You will get good at recognizing which transformations are linear. If the output involves squares or products or lengths, v_1^2 or v_1v_2 or $\|v\|$, then T is *not linear.*

Notes

Exercise

✔ Problem. Determine if the length $T(\vec{v}) = \|v\|$ is linear or not.

Solution.

This T is not linear. Requirement (1) for linearity is $\|v+w\| = \|v\| + \|w\|$. As we know, the sides of a triangle satisfy the *triangle inequality* $\|v+w\| \le \|v\| + \|w\|$.

Requirement (2) would be $\|cv\| = c\|v\|$. The length $\|-v\|$ is not $-\|v\|$. For negative c, it fails.

Notes

Range and Kernel of T

✔ Recall the column space that consisted of all outputs $A\vec{v}$, and the nullspace that consisted of all inputs for which $A\vec{v} = 0$.

✔ Where there is no matrix, we can't talk about a column space. But the idea can be rescued for transformations, and translate those into *range* and *kernel*:

- Range of T = set of all outputs $T(\vec{v})$ (similar to the concept of column space).
- Kernel of T = set of all inputs for which $T(\vec{v}) = \vec{0}$ (similar to the concept of nullspace).

✔ The range is in the output vector space $\mathcal{W}$. The kernel is in the input vector space $\mathcal{V}$. When the transformation T is the matrix multiplication $T(\vec{v}) = A\vec{v}$, we can express range and kernel as column space and nullspace, respectively.

Notes

Rotations

✔ Problem. Let T be the transformation that *rotates every vector by* $30°$. The "*domain*" is the xy plane (all input vectors $\vec{v}$). The "*range*" is also the xy plane ($T(\vec{v})$ are all rotated vectors). We describe T without a matrix: rotate by $30°$.

Determine if the rotation is linear.

Solution.

This T is linear. We can rotate two vectors and add the results. The sum of rotations $T(\vec{v}) + T(\vec{w})$ is the same as the rotation $T(\vec{v} + \vec{w})$ of the sum. The whole plane is turning together in this linear transformation.

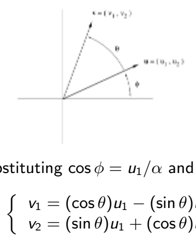
One way to show that T is linear is to demonstrate that $T(\vec{v}) = R\vec{v}$, where R is a suitable matrix. This is done in the next problem.

Notes

Rotations

✔ Problem. Consider the transformation T that rotates a vector of $\mathbb{R}^2$ by the angle θ. Here $\mathcal{V} = \mathcal{W} = \mathbb{R}^2$. Find a matrix A for T.

Solution. If a nonzero vector $\vec{u} = (u_1, u_2)$ is rotated counterclockwise through an angle θ to produce $\vec{v} = (v_1, v_2)$, how are the coordinates of $\vec{v}$ related to $\vec{u}$?



$$v_1 = \alpha\cos(\phi + \theta) = \alpha(\cos\theta\cos\phi - \sin\theta\sin\phi),$$
$$v_2 = \alpha\sin(\phi + \theta) = \alpha(\sin\theta\cos\phi + \cos\theta\sin\phi).$$

where $\alpha = \|u\| = \|v\|$.

Substituting $\cos\phi = u_1/\alpha$ and $\sin\phi = u_2/\alpha$ yields

$$\begin{cases} v_1 = (\cos\theta)u_1 - (\sin\theta)u_2, \\ v_2 = (\sin\theta)u_1 + (\cos\theta)u_2, \end{cases} \iff \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \underbrace{\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}}_{R} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Notes

Rotations

In other words, $\vec{v} = R\vec{u}$, where R is the rotator (or rotation operator)

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

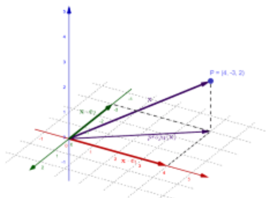
Notice that $R^TR = I$ (R is an orthogonal matrix). If $\vec{v} = R\vec{u}$, then $\vec{u} = R^T\vec{v}$. R^T is a rotator in the opposite direction of that associated with R.

R^T is the rotator associated with the angle $-\theta$. This is confirmed by the fact that if θ is replaced by $-\theta$, then R^T is produced.

Notes

Projections

- ✔ Project every 3-dimensional vector straight down onto the xy plane. Then $T(x, y, z) = (x, y, 0)$. The range is that plane, which contains every $T(\vec{v})$. The kernel is the z axis (which projects down to zero). *This projection is linear.*



Note. Project every 3-dimensional vector onto the horizontal plane $z = 1$. The vector $\vec{v} = (x, y, z)$ is transformed to $T(\vec{v}) = (x, y, 1)$. *This projection is not linear.* Why not ? It doesn't even transform $\vec{v} = 0$ into $T(\vec{v}) = \vec{0}$. □

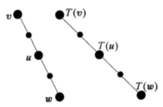
Notes

Lines to Lines

- ✔ The figure below shows the line from $\vec{v}$ to $\vec{w}$ in the input space. It also shows the line from $T(\vec{v})$ to $T(\vec{w})$ in the output space.

Linearity tells us that:

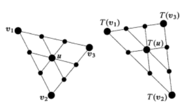
1. *every point on the input line is transformed to a point on the output line*;
2. *equally spaced points go to equally spaced points*. The middle point $\vec{u} = \frac{1}{2}\vec{v} + \frac{1}{2}\vec{w}$ goes to the middle point $T(\vec{u}) = \frac{1}{2}T(\vec{v}) + \frac{1}{2}T(\vec{w})$.



Notes

Triangles to Triangles

- ✔ The figure below moves up a dimension.



Now we have three corners $\vec{v}_1, \vec{v}_2, \vec{v}_3$.

Those inputs have three outputs $T(\vec{v}_1), T(\vec{v}_2), T(\vec{v}_3)$. The input triangle goes onto the output triangle.

- ✔ *Equally spaced points stay equally spaced* (along the edges, and then between the edges). *The middle point or centroid* $\vec{u} = \frac{1}{3}(\vec{v}_1 + \vec{v}_2 + \vec{v}_3)$ *goes to the middle point* $T(\vec{u}) = \frac{1}{3}(T(\vec{v}_1) + T(\vec{v}_2) + T(\vec{v}_3))$.

Notes

Linear Transformations of the Plane

- It is often more interesting to *see* a transformation than to define it.

  Start with a "house" that has eleven endpoints. The columns of H are the eleven corners of the first house.

$$H = \begin{bmatrix} -6 & -6 & -7 & 0 & 7 & 6 & 6 & -3 & -3 & 0 & 0 & -6 \\ -7 & 2 & 1 & 8 & 1 & 2 & -7 & -7 & -2 & -2 & -7 & -7 \end{bmatrix}.$$

  The eleven points in the house matrix H are multiplied by a 2×2 matrix A to produce the corners AH of the other houses. We can watch how it acts.

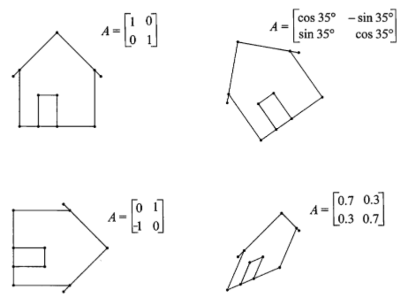
  Straight lines between $\vec{v}$'s become straight lines between the transformed vectors $A\vec{v}$ (the transformation from house to house is linear!)

  Applying A to a standard house produces a new house - possibly stretched or rotated or otherwise unlivable.

Notes

Linear Transformations of the Plane

House matrix $H = \begin{bmatrix} -6 & -6 & -7 & 0 & 7 & 6 & 6 & -3 & -3 & 0 & 0 & -6 \\ -7 & 2 & 1 & 8 & 1 & 2 & -7 & -7 & -2 & -2 & -7 & -7 \end{bmatrix}.$



Notes

Important Question

- We are reaching an important question:

  "can all linear transformations be represented by matrices (of finite size) ?"

- In general, the answer is *"no"*. For example, the differential and integral operators in our previous examples do not have matrix representations because they are defined on infinite-dimensional spaces.
- However, all linear transformations on finite-dimensional spaces, for example from $\mathcal{V} = \mathbb{R}^n$ to $\mathcal{W} = \mathbb{R}^m$, are produced by matrices. When a linear T is described as a "rotation" or "projection", there is always a matrix hiding behind T.
- To see why, the concept of "coordinates" in higher dimensions must first be understood.

Notes

Outline

Linear Operators

The Matrix of a Transformation

Composite Transformations

Notes

The Matrix of a Linear Transformation

- We show how to assign a matrix to every linear transformation $T : \mathcal{V} \to \mathcal{W}$, with $\dim(\mathcal{V}) = n$, $\dim(\mathcal{W}) = m$. *The matrix A for this transformation T will be $m \times n$.*
- We choose a basis $\vec{v}_1, \dots, \vec{v}_n$ for $\mathcal{V}$ and a basis $\vec{w}_1, \dots, \vec{w}_m$ for $\mathcal{W}$. Our choice of bases in $\mathcal{V}$ and $\mathcal{W}$ will decide A.
- As in general the basis of a vector space is not unique, the same T may be represented by different matrices.
- Key idea: compute $T(\vec{v}_1), \dots, T(\vec{v}_n)$ for the basis vectors $\vec{v}_1, \dots \vec{v}_n$ of $\mathcal{V}$. Then linearity produces $T(\vec{v})$ for every other input vector $\vec{v}$. The output $T(\vec{v}_1)$ is in $\mathcal{W}$.

  $T(\vec{v}_1)$ is a combination $a_{11}\vec{w}_1 + \dots + a_{m1}\vec{w}_m$ of the output basis for $\mathcal{W}$.

  These numbers $a_{11}, \dots, a_{m1}$ go into the first column of A.

Notes

Construction of the Matrix

- Repeat the same for the other columns of A. Here is the construction of A.

  Key rule: *The j-th column of A is found by applying T to the j-th basis vector $\vec{v}_j$.*

  $T(\vec{v}_j)$ is a combination of basis vectors of $\mathcal{W}$: $a_{1j}\vec{w}_1 + \dots + a_{mj}\vec{w}_m$.

  These numbers $a_{1j}, \dots, a_{mj}$ go into the j-th column of A.
- At the end of the process, you will get the following coordinate matrix:

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} = \begin{bmatrix} \vec{a}_1 & \vec{a}_2 & \cdots & \vec{a}_n \end{bmatrix}$$

  where $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ are the columns of A. Why is A the correct matrix ?

Notes

Construction of the Matrix

- $T(\vec{v}_j)$ is a combination of basis vectors of $\mathcal{W}$: $a_{1j}\vec{w}_1 + \dots + a_{mj}\vec{w}_m$.
- Every $\vec{v}$ is a combination $c_1\vec{v}_1 + \dots + c_n\vec{v}_n$:

$$\vec{v} = c_1\vec{v}_1 + \dots + c_n\vec{v}_n,$$

  and by linearity

$$T(\vec{v}) = c_1 T(\vec{v}_1) + \dots + c_n T(\vec{v}_n) = c_1 \sum_{i=1}^{m} a_{i1}\vec{w}_i + \dots + c_n \sum_{i=1}^{m} a_{in}\vec{w}_i =$$
$$\left(\sum_{j=1}^{n} c_j a_{1j}\right)\vec{w}_1 + \dots + \left(\sum_{j=1}^{n} c_j a_{mj}\right)\vec{w}_m = (A\vec{c})_1\,\vec{w}_1 + \dots + (A\vec{c})_m\,\vec{w}_m$$

  with $\vec{c} = (c_1, \dots, c_n)$ and $(A\vec{c})_i$ denotes the i-th element of the vector $A\vec{c}$.

Notes

Construction of the Matrix

- When A multiplies the coefficient vector $\vec{c} = (c_1, \dots, c_n)$ in the $\vec{v}$ combination, $A\vec{c}$ produces the coefficients in the $T(\vec{v})$ combination of the basis vectors $\vec{w}_1, \dots, \vec{w}_m$ for $\mathcal{W}$. This is because matrix multiplication (combining columns) is linear like T.
- The discussion made three points.
  1. *The coordinate matrix A tells us what T does.*
  2. *Every linear transformation from an n-dimensional space $\mathcal{V}$ to an m-dimensional space $\mathcal{W}$ can be represented by an $m \times n$ matrix.*
  3. *The coordinate matrix A depends on the bases.*

Notes

Example 1

✔ Problem. Suppose T transforms $\vec{v}_1 = (1,0)$ to $T(\vec{v}_1) = (2,3,4)$ while the second basis vector $\vec{v}_2 = (0,1)$ goes to $T(\vec{v}_2) = (5,5,5)$. Compute $T(\vec{v})$ for $\vec{v} = (1,1)$.

Solution. $T(1,1) = T((1,0)+(0,1)) = T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = (2,3,4) + (5,5,5) = (7,8,9)$. We use the standard basis for $\mathbb{R}^3$.

Those outputs $T(\vec{v}_1)$ and $T(\vec{v}_2)$ go into the columns of a 3×2 matrix A:

$$A = [T(\vec{v}_1), T(\vec{v}_2)] = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix},$$

and we can write the operation $T(\vec{v})$ as a matrix-vector multiplication

$$T(\vec{v}) = T(\vec{v}_1 + \vec{v}_2) = T(\vec{v}_1) + T(\vec{v}_2) = A\vec{v} = \begin{bmatrix} 2 & 5 \\ 3 & 5 \\ 4 & 5 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix}.$$

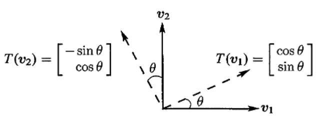
Note that if T is linear from $\mathbb{R}^2$ to $\mathbb{R}^3$ then its "standard matrix" is 3 by 2.

Notes

Example 2 *(Rotation)*

✔ Problem. T rotates every vector by an angle θ. Here $\mathcal{V} = \mathcal{W} = \mathbb{R}^2$. Find matrix A.

Solution. The standard basis is $\vec{v}_1 = (1,0)$ and $\vec{v}_2 = (0,1)$. We use the same basis for $\mathcal{W}$. To find A, apply T to $\vec{v}_1, \vec{v}_1$. In the figure below, they are rotated by θ.



The first vector $(1,0)$ swings around to $(\cos\theta, \sin\theta)$. Therefore those numbers $\cos\theta$ and $\sin\theta$ go into the first column of A.

Notes

Example 2 *(Rotation)*

The figure shows the second vector $(0,1)$ rotated to $(-\sin\theta, \cos\theta)$. Those numbers go into the second column.

$$\begin{bmatrix} \cos\theta & \\ \sin\theta & \end{bmatrix} \text{ shows column 1} \quad A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \text{ shows both columns.}$$

Multiplying A times $(1,0)$ produces $(\cos\theta, \sin\theta)$, the first column. Multiplying A times $(0,1)$ produces $(-\sin\theta, \cos\theta)$, the second column. A agrees with T on the basis, and on all vectors $\vec{v}$. Hence we found the matrix associated to the linear transformation *rotation*.

Notes

Example 3 *(Projection)*

✔ Problem. Suppose T is the transformation that projects every vector of $\mathbb{R}^3$ onto the xy plane. Find its matrix. Use the standard basis for $\mathbb{R}^3$.

Solution. $T(\vec{u}) = T(u_1, u_2, u_3) = (u_1, u_2, 0)$. We will find the coordinate matrix:

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{as} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix}.$$

Once again, the transformtaion T can be represented by a matrix.

Notes

Example 3 *(Identity Transformation)*

- **Problem.** Find the matrix for the linear transformation $T(\vec{v}) = \vec{v}$ from $\mathbb{R}^m$ to $\mathbb{R}^m$ and find A, using the standard basis for $\mathbb{R}^m$.

  Solution. This identity transformation does nothing to $\vec{v}$.

  The output $T(\vec{v}_1)$ is $\vec{v}_1$. When the bases are the same, this is $\vec{w}_1$. So the first column of A is $(1, 0, ..., 0)$. When each output $T(\vec{v}_j) = \vec{v}_j$ is the same as $\vec{w}_j$, the matrix is just I.

  The identity transformation is represented by the identity matrix.

Notes

Example 3 *(Identity Transformation)*

- But if the bases are *different*, then $T(\vec{v}_1) = \vec{v}_1$ is a combination of the $\vec{w}_i$'s. That combination $m_{11}\vec{w}_1 + \ldots + m_{n1}\vec{w}_n$ tells the first column of the matrix (call it M).
- When the outputs $T(\vec{v}_j) = \vec{v}_j$ are combinations $\sum_{i=1}^{n} m_{ij}\vec{w}_i$ the "*change of basis matrix*" for the identity transformation is M.
- When the inputs have one basis and the outputs have another basis, the matrix is not I. The basis is changing but the vectors themselves are not changing: $T(\vec{v}) = \vec{v}$.
- We examine this case in the next example.

Notes

Example 4 *(Change of basis)*

- **Problem.** Consider the identity transformation $T(\vec{v}) = \vec{v}$. The input basis is $\vec{v}_1 = (3, 7)$ and $\vec{v}_2 = (2, 5)$. The output basis is $\vec{w}_1 = (1, 0)$ and $\vec{w}_2 = (0, 1)$. Compute the change of basis matrix M.

  Solution. If the output bases were equal to the input bases ($\vec{w}_1 = \vec{v}_1$ and $\vec{w}_2 = \vec{v}_2$), then $M = I$. But the output basis is $\vec{w}_1 = (1, 0) \neq \vec{v}_1$ and $\vec{w}_2 = (0, 1) \neq \vec{v}_2$.

  The first input is the basis vector $\vec{v}_1 = (3, 7)$. The output is $T(\vec{v}_1) = \vec{v}_1 = (3, 7)$ which we express as $3\vec{w}_1 + 7\vec{w}_2$. Then the first column of M contains 3 and 7.

  The change of basis matrix for $T(\vec{v}) = \vec{v}$ is easy to compute: $M = \begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}$.

  This seems too simple to be important. It becomes trickier when the change of basis goes the other way. We get the inverse of the previous matrix M.

Notes

Example 4 *(Change of basis)*

Suppose now that the input basis is $\vec{v}_1 = (1, 0)$ and $\vec{v}_2 = (0, 1)$. The outputs are just $T(\vec{v}) = \vec{v}$. But the output basis is now $\vec{w}_1 = (3, 7)$ and $\vec{w}_2 = (2, 5)$.

Reverse the bases
Invert the matrix

The matrix for $T(\vec{v}) = \vec{v}$ is $\begin{bmatrix} 3 & 2 \\ 7 & 5 \end{bmatrix}^{-1} = \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix}$.

Reason. The first input is $\vec{v}_1 = (1, 0)$. The output is also $\vec{v}_1$ but we express it as $m_{11}\vec{w}_1 + m_{21}\vec{w}_2 = 5\vec{w}_1 - 7\vec{w}_2$. Check that $5\,(3, 7) - 7\,(2, 5)$ does produce $(1, 0)$.

We are combining the columns of the previous M to get the columns of I. The matrix to do that is M^{-1}.

Notes

Example 4 *(Change of basis)*

Change basis
Change back $\begin{bmatrix} \vec{w}_1 & \vec{w}_2 \end{bmatrix} \begin{bmatrix} 5 & -2 \\ -7 & 3 \end{bmatrix} = \begin{bmatrix} \vec{v}_1 & \vec{v}_2 \end{bmatrix}$ is $MM^{-1} = I$.

Since we start and end with the same basis $(1,0)$ and $(0,1)$, matrix multiplication must give I. So the two change of basis matrices are inverses.

Multiplying M^{-1} times $(1,0)$ gives column 1 of the matrix M^{-1}. That $(1,0)$ stands for the first vector $\vec{v}_1$, *written in the basis of $\vec{v}$'s*. Then column 1 of the matrix M^{-1} is that same vector $\vec{v}_1$, *written in the standard basis $\vec{w}_1$, $\vec{w}_2$*.

Notes

Outline

Linear Operators

The Matrix of a Transformation

Composite Transformations

Notes

Example *(Projection)*

- ✔ Recall the projection T, that projects every vector of $\mathbb{R}^3$ onto the xy plane: $T(\vec{u}) = T(u_1, u_2, u_3) = (u_1, u_2, 0)$.
- ✔ The coordinate matrix is

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{as} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} = \begin{bmatrix} u_1 \\ u_2 \\ 0 \end{bmatrix}. \quad \square$$

- ✔ Important observation: if you square A, it doesn't change. Projecting twice is the same as projecting once: $T^2 = T$ so $A^2 = A$. Notice what is hidden in that statement: The matrix for T^2 is $AA = A^2$.

Notes

Composite Transformations

- ✔ We have come discovered the real reason for the way matrices are multiplied.
- ✔ Consider two transformations S and T represented by two matrices B and A.
- ✔ The transformation S is from a space $\mathcal{U}$ to $\mathcal{V}$. Its matrix B uses a basis $\vec{u}_1, \ldots, \vec{u}_p$ for $\mathcal{U}$ and a basis $\vec{v}_1, \ldots, \vec{v}_n$ for $\mathcal{V}$. The matrix B is n by p. The transformation T is from $\mathcal{V}$ to $\mathcal{W}$. Its matrix A must use the same basis $\vec{v}_1, \ldots, \vec{v}_n$ for $\mathcal{V}$ - this is the output space for S and the input space for T. Then the matrix AB matches TS.
- ✔ When B multiplies the coefficient vector $\vec{c} = (c_1, \ldots, c_n)$ in the $\vec{u}$ combination, $B\vec{c}$ produces the coefficients in the $S(\vec{u})$ combination of the basis vectors $\vec{v}_1, \ldots, \vec{v}_n$ for $\mathcal{V}$. Then $A(B\vec{c}) = (AB)\vec{c}$ produces the coefficients in the $T(S(\vec{u}))$ combination of the basis vectors $\vec{w}_1, \ldots, \vec{w}_m$ for $\mathcal{W}$.
- ✔ Matrix multiplication gives the correct matrix AB to represent TS.

Notes

Composite Transformations

- The linear transformation TS starts with any vector $\vec{u}$ in $\mathcal{U}$, goes to $S(\vec{u})$ in $\mathcal{V}$ and then to $T(S(\mathcal{U}))$ in $\mathcal{W}$.
- The matrix AB starts with any $\vec{x}$ in $\mathbb{R}^p$, goes to $B\vec{x}$ in $\mathbb{R}^n$ and then to $A(B\vec{x}) = (AB)\vec{x}$ in $\mathbb{R}^m$.
- Then the matrix AB correctly represents TS:

$$TS : \mathcal{U} \to \mathcal{V} \to \mathcal{W} \quad AB : (m \text{ by } n)(n \text{ by } p) = (m \text{ by } p).$$

- Product of transformations matches product of matrices.
- The most important cases are when the spaces $\mathcal{U}$, $\mathcal{V}$, $\mathcal{W}$ are the same and their bases are the same. With $m = n = p$ we have square matrices.

Notes

Example *(Rotation)*

- S rotates vectors by θ and T also rotates by θ. Then TS rotates by 2θ. This transformation $T^2 = TS$ corresponds to the rotation matrix A^2 through 2θ:

$$T = S, \quad A = B, \quad T^2 = \text{rotation by } 2\theta, \quad A^2 = \begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix}.$$

- By the *double angle rule* ($\cos 2\theta = \cos^2\theta - \sin^2\theta$ and $\sin 2\theta = 2\sin\theta\cos\theta$):

$$\begin{bmatrix} \cos 2\theta & -\sin 2\theta \\ \sin 2\theta & \cos 2\theta \end{bmatrix} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}.$$

Notes

Example *(Rotation)*

- S rotates by θ and T rotates by $-\theta$. Then $TS = I$ matches $AB = I$.

  In this case $T(S(\vec{u}))$ is $\vec{u}$. We rotate forward and back. For the matrices to match, $A(B\vec{x}) = (AB)\vec{x}$ must be x. The two matrices are inverses.

  Check this by putting $\cos(-\theta) = \cos\theta$ and $\sin(-\theta) = -\sin\theta$ into the backward rotation matrix:

$$AB = \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$

$$= \begin{bmatrix} \cos^2\theta + \sin^2\theta & 0 \\ 0 & \cos^2\theta + \sin^2\theta \end{bmatrix} = I$$

Notes

Notes