ECON8000: Quantitative Skills for Economics
Lecture 5: Linear Algebra

Shino Takayama

University of Queensland

February 2020

Linear Transformations

If V and W are vector spaces over F, a function $L : V \to W$ is a linear transformation if:

- $L(u + v) = L(u) + L(v)$ for all $u, v \in V$.
- $L(\alpha v) = \alpha L(v)$ for all $\alpha \in F$ and $v \in V$.

The identity function on V is a linear transformation.

If $M : W \to X$ is a second linear transformation, then $M \circ L$ is a linear transformation because

$$M(L(u + v)) = M(L(u) + L(v)) = M(L(u)) + M(L(v))$$

and

$$M(L(\alpha v)) = M(\alpha L(v)) = \alpha M(L(v)).$$

Inverse Linear Transformations

Theorem: If $L : V \to W$ is a linear transformation and a bijection, then L^{-1} is a linear transformation.

Proof. Fix $u, v \in W$ and $\alpha \in F$. Then

$$L^{-1}(u+v) = L^{-1}(L(L^{-1}(u)) + L(L^{-1}(v)))$$
$$= L^{-1}(L(L^{-1}(u) + L^{-1}(v))) = L^{-1}(u) + L^{-1}(v)$$

and

$$L^{-1}(\alpha v) = L^{-1}(\alpha L(L^{-1}(v))) = L^{-1}(L(\alpha L^{-1}(v))) = \alpha L^{-1}(v).$$

Linear Independence, Spans, and Bases

There are three key concepts related to a vector space V:

- A set $S \subset V$ is linearly (in)dependent if there (do not) exist distinct $v_1, \dots, v_k \in S$ and scalars $\alpha_1, \dots, \alpha_k \in F$, not all of which are 0, such that $\alpha_1 v_1 + \cdots + \alpha_k v_k = 0$.
- The span of a set $S \subset V$ is the set of all linear combinations $\alpha_1 v_1 + \cdots + \alpha_k v_k$, where $v_1, \dots, v_k \in S$ and $\alpha_1, \dots, \alpha_k \in F$.
- A set $B \subset V$ is a basis of V if it is linearly independent and its span is all of V.

The Matrix of a Linear Transformation

If $L : V \to W$ is a linear transformation and $v_1, \ldots, v_n$ and $w_1, \ldots, w_m$ are bases for V and W, then there are scalars a_{ij} such that $L(v_j) = a_{1j}w_1 + \cdots + a_{mj}w_m$ for each $j = 1, \ldots, n$. The matrix of L with respect to these bases is

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix}.$$

If $v = d_1v_1 + \cdots + d_nv_n$, then $L(v) = \sum_j d_j(\sum_i a_{ij}w_i) = \sum_i(\sum_j a_{ij}d_j)w_i$. so to compute the coefficient $\sum_j a_{ij}d_j$ of w_i in $L(v)$ we pick up the i^{th} row $\begin{bmatrix} a_{i1} & \cdots & a_{in} \end{bmatrix}$ of A, turn it 90 degrees, and drop it on the column vector $\begin{bmatrix} d_1 \\ \vdots \\ d_n \end{bmatrix}$.

The Transpose

For the $m \times n$ matrix

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \quad \text{the transpose is} \quad A^T = \begin{bmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{mn} \end{bmatrix}.$$

That is, A^T is the $n \times m$ matrix whose ij-entry is a_{ji}.

Matrix Multiplication

If $B = (b_{hi})$ is a second $\ell \times m$ matrix, the product of B and A is

$$BA = \begin{bmatrix} b_{11}a_{11} + \cdots + b_{1m}a_{m1} & \cdots & b_{11}a_{1n} + \cdots + b_{1m}a_{mn} \\ \vdots & & \vdots \\ b_{\ell 1}a_{11} + \cdots + b_{\ell m}a_{m1} & \cdots & b_{\ell m}a_{1n} + \cdots + b_{\ell m}a_{mn} \end{bmatrix}.$$

Permutations

A permutation of a finite set is a bijection from that set to itself. The symmetric group on n elements, denoted by Σ_n, is the set of permutations $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. For any two permutations σ and τ, $\tau \circ \sigma$ is also a permutation.

Swaps

If $1 \leq i < j \leq n$, let $(i\ j)$ be the element of Σ_n such that

$$(i\ j)(k) = \begin{cases} i, & k = j, \\ j, & k = i, \\ k, & \text{otherwise.} \end{cases}$$

In standard terminology a permutation such as $(i\ j)$ is called a 2-cycle (you can probably guess what a 3-cycle is) but my preferred term is swap. We let $(i\ i)$ denote the identity.

The Sign of a Permutation

The sign of σ, denoted by $\mathrm{sgn}(\sigma)$ is -1 or 1 according to whether σ is a composition of an odd or even number of swaps. Note that

$$\mathrm{sgn}(\tau \circ \sigma) = \mathrm{sgn}(\tau) \cdot \mathrm{sgn}(\sigma).$$

From this it follows that

$$\mathrm{sgn}(\sigma^{-1}) = \mathrm{sgn}(\sigma).$$

The Determinant

The determinant of an $n \times n$ matrix $A = (a_{ji})$ is

$$\det(A) = \sum_{\sigma \in \Sigma_n} \mathrm{sgn}(\sigma) a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Theorem: If A and B are $n \times n$ matrices, then

$$\det(BA) = \det(A) \cdot \det(B).$$

Proof. Combining the definitions of matrix multiplication and the determinant, we compute that

$$\det(BA) = \sum_{\sigma\in\Sigma_n} \mathrm{sgn}(\sigma)\cdot\Big(\sum_{j=1}^{n} b_{\sigma(1)j}a_{j1}\Big)\times\cdots\times\Big(\sum_{j=1}^{n} b_{\sigma(n)j}a_{jn}\Big)$$

$$= \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n}\sum_{\sigma\in\Sigma_n} \mathrm{sgn}(\sigma) b_{\sigma(1)j_1}a_{j_1 1}\cdots b_{\sigma(n)j_n}a_{j_n n}$$

$$= \sum_{j_1=1}^{n}\cdots\sum_{j_n=1}^{n}\det(C(j_1,\ldots,j_n))$$

where $C(j_1,\ldots,j_n)$ is the matrix with entries $b_{kj_i}a_{j_i i}$. If $j_i = j_{i'}$ for distinct i and i', then this matrix has two identical rows and its determinant vanishes. Therefore we can sum over those $j_1,\ldots,j_n$ that are all different, and it we let τ be the permutation $1\to j_1,\ldots,n\to j_n$ to obtain

$$
\begin{aligned}
\det(BA) &= \sum_{\tau\in\Sigma_n}\sum_{\sigma\in\Sigma_n} \mathrm{sgn}(\sigma) b_{\sigma(1)\tau(1)} a_{\tau(1)1} \cdots b_{\sigma(n)\tau(n)} a_{\tau(n)n} \\
&= \sum_{\theta\in\Sigma_n}\sum_{\sigma\in\Sigma_n} \mathrm{sgn}(\sigma) b_{\sigma(\theta(1))1} a_{1\theta(1)} \cdots b_{\sigma(\theta(n))n} a_{n\theta(n)} \\
&= \sum_{\theta\in\Sigma_n}\sum_{\rho\in\Sigma_n} \mathrm{sgn}(\rho\circ\theta^{-1}) b_{\rho(1)1} a_{1\theta(1)} \cdots b_{\rho(n)n} a_{n\theta(n)} \\
&= \sum_{\theta\in\Sigma_n}\sum_{\rho\in\Sigma_n} \mathrm{sgn}(\rho)\cdot\mathrm{sgn}(\theta) b_{\rho(1)1} a_{1\theta(1)} \cdots b_{\rho(n)n} a_{n\theta(n)} \\
&= \Big(\sum_{\rho\in\Sigma_n} \mathrm{sgn}(\rho) b_{\rho(1)1}\cdots b_{\rho(n)n}\Big)\cdot\Big(\sum_{\tau\in\Sigma_n} \mathrm{sgn}(\tau) a_{\tau(1)1}\cdots a_{\tau(n)n}\Big) \\
&= \det(B)\cdot\det(A).
\end{aligned}
$$

In this calculation we let $\theta = \tau^{-1}$ (note that $\sigma \mapsto \sigma^{-1}$ is a bijection) then set $\rho = \sigma\circ\theta^{-1}$ (note that $\sigma \mapsto \sigma\circ\theta^{-1}$ is a bijection).

Nonsingularity and Invertibility

Let V and W be n dimensional vector spaces, let $L : V \to W$ be a linear transformation, and let $A = (a_{ij})$ be the matrix of L with respect to bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$. We say that L and A are nonsingular if $\det(A) \neq 0$; otherwise they are singular.

The $n \times n$ identity matrix is

$$I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}.$$

An $n \times n$ matrix A is left (right) invertible if there is an $n \times n$ matrix B such that $BA = I_n$ ($AB = I_n$) which we call a left (right) inverse.

Example

Consider

$$X = \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix}.$$

- $(1,2,3) \to (1,2,3)$
- $(1,2,3) \to (1,3,2)$: $2 \leftrightharpoons 3$
- $(1,2,3) \to (2,1,3)$: $1 \leftrightharpoons 2$
- $(1,2,3) \to (2,3,1)$: $1 \leftrightharpoons 2$ and $3 \leftrightharpoons 1$
- $(1,2,3) \to (3,1,2)$: $3 \leftrightharpoons 1$ and $1 \leftrightharpoons 2$
- $(1,2,3) \to (3,2,1)$: $1 \leftrightharpoons 3$

Thus the determinant of X is

$$x_{11}x_{22}x_{33} - x_{11}x_{32}x_{23} - x_{21}x_{12}x_{33} + x_{21}x_{32}x_{13} + x_{31}x_{12}x_{23} - x_{31}x_{22}x_{13}.$$

Observation

Indeed,

$$
\begin{aligned}
&det \begin{bmatrix} x_{11} & x_{12} & x_{13} \\ x_{21} & x_{22} & x_{23} \\ x_{31} & x_{32} & x_{33} \end{bmatrix} \\
&= x_{11} \cdot det \begin{bmatrix} x_{22} & x_{23} \\ x_{32} & x_{33} \end{bmatrix} - x_{21} \cdot det \begin{bmatrix} x_{12} & x_{13} \\ x_{32} & x_{33} \end{bmatrix} \\
&+ x_{31} \cdot det \begin{bmatrix} x_{12} & x_{13} \\ x_{22} & x_{23} \end{bmatrix}.
\end{aligned}
$$

Useful Tip

Let A be an $n \times n$ matrix such that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}.$$

Then

$$det(A) = det(A_{11}) \cdot det(A_{22}).$$

Law 1

$$det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ ca_{j1} & \cdots & ca_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = c \cdot det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}.$$

Law 2

$$
det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{j1} + b_{j1} & \cdots & a_{jn} + b_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}
$$

$$
= det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} + det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ b_{j1} & \cdots & b_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} .
$$

Law 3

$$det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = 0.$$

In the end

$$det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots \\ ca_{k1} + da_{j1} & \cdots & ca_{kn} + da_{jn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = c \cdot det \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \vdots & \vdots \\ a_{j1} & \cdots & a_{jn} \\ \vdots & \vdots & \vdots \\ a_{k1} & \cdots & a_{kn} \\ \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

Solving Systems of Equations

Suppose that V and W are n dimensional vector spaces, A is the matrix of the linear transformation $L : V \to W$ with respect to the bases $v_1, \ldots, v_n$ and $w_1, \ldots, w_n$, and $w = \beta_1 w_1 + \cdots + \beta_n w_n$ is a point in W. Typically we imagine solving the equation $L(v) = w$, which means finding all $v = \alpha_1 v_1 + \cdots + \alpha_n v_n$ that satisfy it. We write this as a system of equations:

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_n \end{bmatrix}.$$

The standard method for solving this is called Gaussian elimination in honor of Carl Friedrich Gauss, even though it appears in a Chinese book from 179 A.D., and was also described by Newton. The idea is to use to first equation to solve for α_1, then substitute this into the other equations, reducing to a system of $n-1$ equations in $n-1$ unknowns.

We can describe this in terms of row operations on the matrix equation above. Supposing that $a_{11} \neq 0$, we add $-a_{n1}/a_{11}$ times the first equation, to the last equation, obtaining

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & \vdots & & \vdots \\ a_{n-1,1} & a_{n-1,2} & \cdots & a_{n-1,n} \\ 0 & a_{n2} - \frac{a_{n1}}{a_{11}} a_{12} & \cdots & a_{nn} - \frac{a_{n1}}{a_{11}} a_{1n} \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} \beta_1 \\ \vdots \\ \beta_{n-1} \\ \beta_n - \frac{a_{n1}}{a_{11}} \beta_1 \end{bmatrix}.$$

One can repeat this until all entries in the first column, except for a_{11}, are zero. The second equation is then used to eliminate α_2 from all equations except the second, and so forth. Eventually we obtain a diagonal matrix, and at that point we multiply the i^{th} equation by $1/a_{ii}$.

An Example

We solve the system

$$\begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 2 \\ -2 & 1 & 2 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -11 \\ -3 \end{bmatrix}.$$

Successive row operations yield

$$\begin{bmatrix} 2 & 1 & -1 \\ -3 & -1 & 2 \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 8 \\ -11 \\ 5 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 1 & -1 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 8 \\ 1 \\ 5 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & -2 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 2 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 5 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & -2 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 6 \\ 1 \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & -2 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 6 \\ \frac{3}{2} \\ 1 \end{bmatrix},$$

$$\begin{bmatrix} 2 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 4 \\ \frac{3}{2} \\ 1 \end{bmatrix},$$

and

$$\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ -1 \end{bmatrix}.$$

Obviously the column vector we are solving for doesn't change, and doesn't add any information, so when done by hand the algorithm is usually expressed in terms of an augmented matrix in which the last column is what for us is the right hand side.