Basic Linear Algebra

Ichiro Obara

UCLA

September 12, 2019

Basic Linear Algebra in $\Re^n$

We are familiar with $\Re$ (real line), $\Re^2$ (plane), $\Re^3$ (space). In Economics, we need to deal with spaces in "higher dimension". We would like to extend what we know and our intuition in these spaces to higher dimensional spaces.

It will be helpful to try to understand all the concepts and the results in the above basic spaces, for which we have a good geometric intuition.

Vector

A **vector** in $\Re^n$ is an n-tuple of numbers. There are two expressions of vectors:

$\mathbf{x} = (x_1, ..., x_n)$ is a **row vector** and $\mathbf{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ is a **column vector**. We will use row vectors in the following to save some space.

Addition/subtraction/scalar multiplication can be naturally defined as follows.

- **Addition:** $\mathbf{x} + \mathbf{y} = (x_1 + y_1, ..., x_n + y_n)$.
- **Subtraction:** $\mathbf{x} - \mathbf{y} = (x_1 - y_1, ..., x_n - y_n)$.
- **Multiplication:** $a\mathbf{x} = (ax_1, ..., ax_n)$ for $a \in \Re$.

Length

The **length** of a vector $\mathbf{x}$ is given by $\|\mathbf{x}\| = \sqrt{\sum_{i=1}^{n} x_i^2}$.

For any two points $\mathbf{x}, \mathbf{y}$ in $\Re^n$, $\mathbf{x} - \mathbf{y}$ is a vector that connects these two points.

$\|\mathbf{x} - \mathbf{y}\|$ corresponds to the distance between $\mathbf{x}$ and $\mathbf{y}$.

Inner Product

Inner product of two vectors $\mathbf{x}, \mathbf{y} \in \Re^n$ is defined by $\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i$.

Inner product is related to length: $\mathbf{x} \cdot \mathbf{x} = \|\mathbf{x}\|^2$.

The inner product of two orthogonal vectors $(1, -1)$ and $(1, 1)$ is 0. More generally, inner product is related to the angle between two vectors. Consider the following number for nonzero vector $\mathbf{x}$ and $\mathbf{y}$: $\frac{\mathbf{x} \cdot \mathbf{y}}{\|\mathbf{x}\| \|\mathbf{x}\|}$.

- This is 0 if and only if $\mathbf{x}$ and $\mathbf{y}$ are orthogonal.
- This is 1 if and only if $\mathbf{x}$ and $\mathbf{y}$ are parallel (i.e. $\mathbf{y}$ is some scalar multiplication of $\mathbf{x}$).

In fact, this number is $\cos\theta$, where θ is the angle between $\mathbf{x}$ and $\mathbf{y}$.

Hyperplane

An equation $3x_1 + 4x_2 = 12$ represents a **line** in $\Re^2$. Any equation of the form $a_1x_1 + a_2x_2 = b$ is a line in $\Re^2$. Similarly $a_1x_1 + a_2x_2 + a_3x_3 = b$ represents a **plane** in $\Re^3$. In $\Re^n$, the set of $\mathbf{x}$ that satisfies $\sum_{i=1}^{n} a_ix_i = b$, where **normal vector** $\mathbf{a} = (a_1, ..., a_n)$ is not a zero vector, is called **hyperplane**.

Here is a way to get some geometric intuition. Take $a_1x_1 + a_2x_2 = b$ in $\Re_2$. Pick any $\mathbf{x}' \in \Re^2$ such that $a_1x_1' + a_2x_2' = b$. Then $a_1(x_1 - x_1') + a_2(x_2 - x_2') = 0$. We can express this using vector notations and inner product: $\mathbf{a} \cdot (\mathbf{x} - \mathbf{x}') = \mathbf{0}$. Thus the hyperplane is the set of $\mathbf{x}$ such that the normal vector $\mathbf{a}$ and $\mathbf{x} - \mathbf{x}'$ are orthogonal.

Linear Independence

A set of vectors $\mathbf{x}_1, \cdots \mathbf{x}_m$ in $\Re^n$ are **linearly independent** if $\sum_{i=1}^m a_i \mathbf{x}_i = \mathbf{0} \rightarrow a_i = 0$ for $i = 1, \cdots, m$.

we can guess from spaces like $\Re^2$ and $\Re^3$ that the following properties are satisfied.

- In $\Re^n$, the number of linearly independent vectors is at most n.
- If there are m linearly independent vectors $\mathbf{x}_1, \cdots, \mathbf{x}_m$ in $\Re^n$ for $m < n$, then we can find another vector $\mathbf{x}_{m+1}$ that can be added to make $m+1$ linearly independent vectors. In fact we can find $\mathbf{x}_{m+1}$ that is orthogonal to $\mathbf{x}_1, ..., \mathbf{x}_m$ and

Linear Subspace

Linear subspace of $\Re^n$ is a subset of $\Re^n$ that is closed with respect to addition and scalar multiplication. That is, for any $\mathbf{x}, \mathbf{y}$ in linear subspace $X \subset \Re^n$, $\mathbf{x} + \mathbf{y} \in X$ and $a\mathbf{x} \in X$ for any $a \in \Re$.

For example, any straight line that goes through the origin in $\Re^2$ or $\Re^3$ is a linear subspace. As an another example, any hyperplane that goes through the origin in $\Re^n$ is a linear subspace of $\Re^n$.

Spanning Set and Basis

For any set of vectors $\mathbf{x}_1, \cdots, \mathbf{x}_m \in \Re^n$, consider the following set $\left\{x \in \Re^n | \exists a_i, i = 1, ..., m \ s.t. \ x = \sum_{i=1}^m a_i \mathbf{x}_i\right\}$. We say that this set is **spanned by** $\mathbf{x}_1, \cdots, \mathbf{x}_m$. It is easy to check that this set is a linear subspace of $\Re^n$.

For any linear subspace of $\Re^n$, there exists a set of linearly independent vectors in it that spans it. It is called a **basis** of that linear subspace. For example, $\Re^n$ itself is a linear subspace of $\Re^n$. One basis of $\Re^n$ is $\mathbf{e}_i = (0, ..., \underbrace{1}_{ith}, ..., 0), i = 1, ..., n$.

There are many basis for one linear subspace: both $\{(1,0),(0,1)\}$ and $\{(2,1),(1,2)\}$ is a basis of $\Re^2$. However, the number of elements in every basis is the same. This number is the **dimension** of the linear subspace.

For a given basis $\mathbf{x}_1, \cdots, \mathbf{x}_m$, any vector $\mathbf{x}$ in the linear subspace it spans can be expressed uniquely, i.e. there exists unique $a_1, ..., a_m$ such that $\mathbf{x} = \sum_{i=1}^m a_i \mathbf{x}_i$.

Projection and Orthogonalization

Remember that $\cos\theta = \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|\|\mathbf{y}\|}$. So we should expect that $\mathbf{y} - \|\mathbf{y}\|\cos\theta\frac{\mathbf{x}}{\|\mathbf{x}\|} = \mathbf{y} - \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$ is orthogonal to $\mathbf{x}$ (Draw a graph). In fact,

$$\mathbf{x}\left(\mathbf{y} - \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}\right) = 0.$$

- Geometrically, $\frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$ is the point on the linear subspace spanned by $\mathbf{x}$ that is closest to $\mathbf{y}$. This point is called the **projection** of $\mathbf{y}$ to this linear subspace.
- In this way, we can generate two orthogonal vectors ($\mathbf{x}$ and $\mathbf{y} - \frac{\mathbf{x}\cdot\mathbf{y}}{\|\mathbf{x}\|^2}\mathbf{x}$) from any two linearly independent vectors $\mathbf{x}, \mathbf{y}$. Note that they would span exactly the same space.

More generally, consider any linear subspace of $\Re^n$ spanned by a set of linearly independent and orthogonal vectors $\mathbf{x}_1, \cdots \mathbf{x}_m$. Take any vector $\mathbf{y} \in \Re^n$. Then the projection of $\mathbf{y}$ to this subspace is given by

$$\sum_{i=1}^{m} \frac{\mathbf{x}_i \cdot \mathbf{y}}{\|\mathbf{x}_i\|^2} \mathbf{x}_i$$

Note that

$$\mathbf{x}_i \cdot \left(\mathbf{y} - \sum_{i=1}^{m} \frac{\mathbf{x}_i \cdot \mathbf{y}}{\|\mathbf{x}_i\|^2} \mathbf{x}_i \right) = \mathbf{x}_i \cdot \mathbf{y} - \mathbf{x}_i \cdot \mathbf{y} = 0$$

Hence $\mathbf{y} - \sum_{i=1}^{m} \frac{\mathbf{x}_i \cdot \mathbf{y}}{\|\mathbf{x}_i\|^2} \mathbf{x}_i$ is orthogonal to every vector in this linear subspace.

Linear Regression as Projection

Suppose that we have T data points $(x_t, t_t), t = 1, ..., T$ of two variables X and Y and would like to find a linear function of X that explains Y best.

Regard $\mathbf{x}$ and $\mathbf{y}$ as T dimensional vectors. We look for a linear model $a\mathbf{1} + b\mathbf{x}$ that is closest to $\mathbf{y}$, which is equivalent to finding the projection of $\mathbf{y}$ to the linear subspace spanned by $\mathbf{1}$ and $\mathbf{x}$.

Note that $\mathbf{1}$ and $\mathbf{x}$ are not orthogonal. But if we instead use $\mathbf{1}$ and $\mathbf{x} - \overline{x}\mathbf{1}$ (deviation from the average $\overline{x} = \frac{\sum_t x_t}{T}$), then they are orthogonal, hence we can apply the previous formula to find the projection in terms of these vectors. It is

$$\frac{\mathbf{1} \cdot \mathbf{y}}{T}\mathbf{1} + \frac{(\mathbf{x} - \overline{x}\mathbf{1}) \cdot \mathbf{y}}{\|(\mathbf{x} - \overline{x}\mathbf{1})\|^2}(\mathbf{x} - \overline{x}\mathbf{1})$$

which can be expressed more simply as $a\mathbf{1} + b\mathbf{x}$ where $a = \overline{y} - b\overline{x}$ and $b = \frac{\sum_t (x_t - \overline{x})(y_t - \overline{y})}{\sum_t (x_t - \overline{x})^2}$. This b is exactly the coefficient for linear regression.

Exercises

1. Prove the **Cauchy-Schwarz** inequality: $\mathbf{x} \cdot \mathbf{y} \leq \|\mathbf{x}\| \, \|\mathbf{y}\|$.

2. $\|\mathbf{x} - \mathbf{y}\|$ can be interpreted as the distance between point $\mathbf{x}$ and $\mathbf{y}$. Then we should expect that $\|\mathbf{x} - \mathbf{y}\|$ is smaller than $\|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\|$ for any $\mathbf{z} \in \Re^n$. Show that this inequality indeed holds.

3. Let $\mathbf{x}_1, ..., \mathbf{x}_m \in \Re^n$ be a basis that spans some linear subspace.
   - When $n > m$, find $\mathbf{x}_{m+1} \neq \mathbf{0}$ that is orthogonal to $\mathbf{x}_1, ..., \mathbf{x}_m$.
   - Can you find an **orthogonal basis** $\hat{\mathbf{x}}_1, ..., \hat{\mathbf{x}}_m$ ($\hat{\mathbf{x}}_i \cdot \hat{\mathbf{x}}_j = 0$) that spans the same linear subspace? How?

4. What is the dimension of a hyperplane in $\Re^n$ that goes through the origin?

Matrix

A **matrix** A is an rectangular array of numbers. If a matrix has m rows and n columns, it is called a $m \times n$ matrix. The ij element of A is denoted by a_{ij}. We sometimes use $\tilde{\mathbf{a}}_i$ as the ith row vector and $\mathbf{a}_j$ as the jth column vector of A.

We will use matrices to express coefficients of linear system of equations, the derivative of multivariable functions, and asset return structures etc...

Matrix Algebra

- Addition, subtraction, scalar multiplication can be defined element-wise as we did for vectors.
- Some two matrices can be multiplied. For any $\ell \times m$ matrix A and $m \times n$ matrix B, AB is defined as the $\ell \times n$ matrix, where its ij element is given by $\sum_{k=1}^{m} a_{ik} b_{kj}$.
  - For matrix multiplication, the distribution laws such as $A(B + C) = AB + AC$ and $(A + B)C = AC + BC$ hold.
  - But $AB = BA$ does not hold.
- For a $m \times n$ matrix A, the **transpose** of A is the $n \times m$ matrix $A^\top$ whose ij element is given by the ji element of A.

Matrix as Linear Transformation

We can think of an $m \times n$ matrix as a mapping from $\Re^n$ to $\Re^m$, because $A\mathbf{x} \in \Re^m$ for any $\mathbf{x} \in \Re^n$.

Furthermore it is a linear mapping in the sense that: $A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$ and $A(a\mathbf{x}) = aA\mathbf{x}$. Thus A can be regarded as a **linear transformation** from $\Re^n$ to $\Re^m$.

Some Special Matrices

- **Square Matrix:** $n \times n$ matrices.
- **Identity Matrix:** Square matrix where $a_{ii} = 1$ for $i = 1, ..., n$ and all the other elements are 0.
- **Symmetric Matrix:** Square matrix where $a_{ij} = a_{ji}$ for all i, j.

Inverse

Square matrix B is the **inverse** of square matrix A if $AB = BA = I$. The inverse of A is denoted by A^{-1}.

For example, the inverse of $\begin{pmatrix} 2 & 1 \\ -2 & 1 \end{pmatrix}$ is $\begin{pmatrix} 0.25 & -0.25 \\ 0.5 & 0.5 \end{pmatrix}$

A square matrix has either no inverse or unique inverse. Some matrix such as $\begin{pmatrix} 0 & 2 \\ 0 & 1 \end{pmatrix}$ clearly does not have any inverse.

Rank

For any matrix A, the number of linearly independent row vectors and the number of linearly independent column vectors are the same. This number is called the **Rank** of A and is denoted by $Rank(A)$.

For $m \times n$ matrix A, consider the following set $\{\mathbf{y} \in \Re^m | \exists \mathbf{x} \in \Re^n \ s.t. \ \mathbf{y} = A\mathbf{x}\}$. This set is a linear subspace of $\Re^m$ and its dimension is $Rank(A)$ as it is the space spanned by the column vectors $\mathbf{a}_j, j = 1, ..., n$.

Determinant

For 1×1 square matrix $A = (a)$, its **determinant** $detA$ is a. The **determinant** of a $n \times n$ square matrix A is defined recursively by

$$detA = \sum_{j=1}^{n} (-1)^{1+j} a_{1j} detC_{1j}(A)$$

where $C_{1j}(A)$ is the $(n-1) \times (n-1)$ matrix that is obtained by deleting the first row vector and the jth column vector from A.

For example,

- $det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc$
- $det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a \begin{pmatrix} e & f \\ h & i \end{pmatrix} - b \begin{pmatrix} d & f \\ g & i \end{pmatrix} + c \begin{pmatrix} d & e \\ g & h \end{pmatrix}$

It is a complicated formula, but once we compute this number, then we can verify the following important properties.

Theorem

For a $n \times n$ square matrix A, the following properties are equivalent.

- $detA \neq 0$.
- $RankA = n$.
- A^{-1} exists.

Negative Definiteness

,

For maximization, the following type of symmetric matrices are important.

- A symmetric matrix $A \in M^n$ is **negative definite** if $\mathbf{x}^\top A\mathbf{x} < 0$ for any $\mathbf{x} \neq 0$ in $\Re^n$.
- A symmetric matrix $A \in M^n$ is **negative semi-definite** if $\mathbf{x}^\top A\mathbf{x} \leq 0$ for any $\mathbf{x}$ in $\Re^n$.

For 1×1 matrix, it is negative definite if and only if its element is strictly negative.

For 2×2 matrix $\begin{pmatrix} a & b \\ b & c \end{pmatrix}$, $a < 0$ and $c < 0$ are necessary for negative definiteness. More generally, $ax_1^2 + 2bx_1x_2 + cx_2^2$ must be strictly negative for any $\mathbf{x} \in \Re^2$. This is equivalent to saying that there is no solution for the quadratic equation about x_1 $ax_1^2 + (b + c)x_1x_2 + cx_2^2 = 0$ for any x_2. The corresponding condition is $b^2 - ac < 0$, which means that the determinant is strictly positive.

More generally, a $n \times n$ symmetric matrix is negative definite if and only if the determinant of the upper $k \times k$ matrices alternate in sign ($-$ for $k = 1$, $+$ for $k = 2$...).

Exercises

1. Compute the determinant of the following matrices.
   - $\begin{pmatrix} -2 & 3 \\ -1 & 3 \end{pmatrix}$
   - $\begin{pmatrix} -1 & 0 & 2 \\ 1 & 3 & -2 \\ 2 & 1 & -4 \end{pmatrix}$
2. The set of $\mathbf{x} \in \Re^n$ such that $A\mathbf{x} = 0$ for $m \times n$ matrix A is called the **null space** of A.
   - Show that this set is a linear subspace.
   - What is the dimension of this space?

Linear System of Equations

Consider a system of m linear equations with n unknowns.

$$
\begin{aligned}
a_{1,1}x_1 + \cdots \quad + a_{1,n}x_n &= b_1 \\
&\vdots \\
a_{m,1}x_1 + \cdots \quad + a_{m,n}x_n &= b_m
\end{aligned}
$$

This system can be compactly written as $A\mathbf{x} = \mathbf{b}$.

When does this system have a solution for any $\mathbf{b}$? When does it have a unique solution?

Note that the first question is the same as asking when the set of column vectors $\mathbf{a}_1, \cdots \mathbf{a}_n$ spans $\Re^m$. So we have the following result.

Theorem

There is a solution for $A\mathbf{x} = \mathbf{b}$ for any $\mathbf{b} \in \Re^m$ if and only if the rank of A is m.

One implication of this is that, if there are more equations than unknowns (i.e. $m > n$), then there is no solution for some b.

The rank of A also matters for the next question. $\mathbf{a}_1, \cdots \mathbf{a}_n$ are not linearly independent if and only if there exists $\mathbf{x}_0 \neq \mathbf{0}$ such that $A\mathbf{x}_0 = \mathbf{0}$. This is equivalent to nonuniqueness of solution (when there is a solution) because (1) if such x_0 exists, then for any solution $\mathbf{x}$, $\mathbf{x} + \mathbf{x}_0$ is a solution as well, and (2) if there are two solutions $\mathbf{x}' \neq \mathbf{x}''$, then $\mathbf{x}_0 = \mathbf{x}' - \mathbf{x}''$ satisfies $A\mathbf{x}_0 = 0$.

Hence we have the following result.

Theorem

Suppose that there is a solution for $A\mathbf{x} = \mathbf{b}$. This solution is unique if and only if the rank of A is n.

One implication of this is that, if there are more unknowns than equations (i.e. $m < n$), then there is either no solution or infinite number of solutions.

Asset Markets

A (financial) **asset** entitles you to receive a certain amount of money in a certain state. Suppose that one of S states realizes in a month from now. Denote the return of asset k at state s by r_{sk}. You need to pay q_k to purchase one unit of asset k today. Let R be the $S \times K$ matrix with r_{sk} as its sk element.

Market with asset $1, ..., K$ is **complete** if any state contingent payoff vector in $\Re^S$ can be generated by some portfolio of assets $\mathbf{z}$. That is, for any $\mathbf{x} \in \Re^S$, there exists $\mathbf{z} \in \Re^K$ such that $\mathbf{x} = R\mathbf{z} = \sum_{k=1}^{K} \mathbf{r}_k z_k$.

By definition the market is complete if and only if the column vectors of R spans $\Re^S$. Hence the market is complete if and only if $RankR = S$. This implies that we need at least S assets make the market complete.

No Arbitrage and Asset Price

Price vector $\mathbf{q}$ is **arbitrage free** if there is no $\mathbf{z}$ such that $\mathbf{q} \cdot \mathbf{z} \leq 0$ and $R\mathbf{z} > 0$. This means that you cannot make money at some state without losing any money today and any money at any state one month later.

Suppose that the market is complete and $\mathbf{q}$ is **arbitrage free**. Introduce any new asset $K+1$ with return vector $\mathbf{r}_{K+1}$ to the market. Since the market is complete, there exists a subset of assets K' such that $\mathbf{r}_{K+1} = \sum_{k' \in K'} a_{k'} \mathbf{r}_{k'}$ for some $a_{k'} \in \Re, k' \in K'$. This implies that the price of this asset q_{K+1} must be $\sum_{k' \in K'} a_{k'} q_{k'}$. Otherwise you can generate lots of money by selling/buying these assets without affecting your return at each state. Then you can use that money to achieve $Rz > 0$

Exercises

1. Show that a system of linear equations $A\mathbf{x} = \mathbf{b}$ has the unique solution if and only if A is a square matrix and $detA \neq 0$. Also show that this solution is given by $A^{-1}\mathbf{b}$.

2. Consider a market with two assets and $S = 2$. Asset 1 is a safe asset, which you can sell at \$10 in every state. Asset 2 is a risky asset, which you can sell at \$15 at state 1 but only at \$5 at state 2. Let (q_1, q_2) be an arbitrage free price vector, say, $(\$10, \$9)$. Now consider the following **option contract**: you receive a unit of asset 2 and are guaranteed to be able to sell it at \$10 in a month from now if you want. What should be the price of this new asset to preserve arbitrage free property?