

Lecture 2: Linear Algebra

1. MATRICES, EIGENVALUES AND EIGENVECTORS

A matrix can be defined as a collection of numbers arranged in a certain way. A matrix can also be viewed as an operator (a linear transformation) from \mathbb{R}^n to \mathbb{R}^n ; but what does a linear transformation defined by a data set mean?

For example, consider a 10×10 matrix whose rows are indexed by companies and columns are indexed by dates. The (i, j) -entry of the matrix is the price of the stock of the i -th company on date j . This is a matrix defined by some real-life data. Can we give a meaning to the transformation defined by this matrix?

We begin by reviewing eigenvalues and eigenvectors of a matrix. We use the notation $\langle v, w \rangle$ for inner product of vectors. It can also be written in matrix multiplication form as $\langle v, w \rangle = v^T \cdot w$ (a vector is to be understood as a column matrix).

Definition 1. Let A be an $n \times n$ matrix of real entries. If a non-zero vector $v \in \mathbb{R}^n$ and a complex number λ satisfies

$$Av = \lambda v,$$

then v is called an *eigenvector* of A and λ is an *eigenvalue* of A .

Does a matrix always have eigenvalues and eigenvectors? To answer this question, note that a real number λ is an eigenvalue of A if and only if

$$(A - \lambda I)v = 0,$$

for some vector v . Hence it follows that $\det(A - \lambda I) = 0$. Note that if one considers λ as a variable, then $\det(A - \lambda I)$ becomes a polynomial of degree n . Therefore we can conclude that there are at least 1 and at most n distinct (complex) eigenvalues of an $n \times n$ matrix.

Note that the matrix A applied to its eigenvector acts as if it is a scalar multiplication. Thus eigenvectors are ‘the important’ directions of a matrix, when considered as an operator.

Definition 2. A matrix A is *diagonalizable* if there exists an invertible matrix P and a diagonal matrix D such that

$$A = P^{-1}DP.$$

Recall that the conjugation operation by an invertible matrix A is equivalent to a change of basis of the underlying space. Hence a matrix is diagonalizable if there exists a basis for which A becomes a diagonal matrix. Hence a matrix A is a diagonalizable matrix if there exists a basis on which A acts as a scaling operator.

2. EIGENVALUES AND EIGENVECTORS OF SYMMETRIC MATRICES

A matrix A is *symmetric* if $A = A^T$. The eigenvalues and eigenvectors of symmetric matrices will be of particular interest to us, since we will encounter many such matrices throughout the course. The following proposition establishes some properties of eigenvalues and eigenvectors of symmetric matrices.

Proposition 3. *Let A be a real symmetric matrix. Then the following holds.*

- (i) *All eigenvalues are real.*
- (ii) *A is orthonormally diagonalizable (i.e., there exists an orthonormal matrix U such that $A = U^{-1}DU$).*

Proof. (i) Let λ be an eigenvalue and v be an eigenvector such that $Av = \lambda v$. Consider

$$\bar{v}^T Av = \bar{v}^T \cdot (\lambda v) = \lambda \|v\|^2.$$

Since $A = \bar{A}^T$, we also have

$$\bar{v}^T Av = \bar{v}^T \bar{A}^T v = (\bar{\lambda} \bar{v}^T) \cdot v = \bar{\lambda} \|v\|^2.$$

Therefore, $\lambda = \bar{\lambda}$, and thus λ is real.

(ii) Let λ_1 be an eigenvalue of A and v_1 be a corresponding eigenvector (recall that all $n \times n$ matrices have at least one eigenvalue). Let

$$W_1 = \{w : \langle w, v_1 \rangle = 0\}$$

be the set of vectors orthogonal to v_1 . Note that W_1 forms a subspace of \mathbb{R}^n . For a vector $w \in W_1$, we have

$$\langle Aw, v_1 \rangle = \langle w, Av_1 \rangle = \lambda \langle w, v_1 \rangle = 0,$$

and therefore W_1 is a vector space of dimension $n - 1$ that is closed under the transformation A .

Let \mathcal{B}' be an orthonormal basis of the vector space W_1 , and note that $\mathcal{B} = \{v_1\} \cup \mathcal{B}'$ forms an orthonormal basis of \mathbb{R}^n . Let U be the vector whose columns are the vectors in the basis $\{v_1\} \cup \mathcal{B}$, where v_1 is the first column. Then the observation above shows that

$$UAU^{-1} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & A_1 \end{bmatrix},$$

where A_1 is a symmetric $(n - 1) \times (n - 1)$ matrix. By repeating the process, we obtain an orthonormal diagonalization of A . \square

The largest eigenvalue of a symmetric matrix also has the following property.

Theorem 4. *If A is symmetric, then $\max_{\|v\|=1} |Av| = \max_{\lambda} |\lambda|$, where the second maximum is taken over all eigenvalues of the matrix.*

3. SINGULAR VALUE DECOMPOSITION

For a diagonalizable matrix A , we refer to $D = UAU^{-1}$ as the *eigenvalue decomposition* of A . We saw that symmetric matrices are diagonalizable. Can we do something similar for general $n \times n$ matrices? What about $m \times n$ matrices? Singular value decomposition is a generalization of the eigenvalue decomposition that achieves this goal. It is a tool which allows one to understand a matrix through a diagonal matrix.

Theorem 5. Every $m \times n$ matrix A_{mn} can be expressed as

$$A_{mn} = U_{mm} \Sigma_{mn} V_{nn}^T$$

for a $m \times m$ orthonormal matrix U_{mm} , a diagonal $m \times n$ matrix Σ_{mn} , and a $n \times n$ orthonormal matrix V_{nn} .

Before proving the theorem, let us first compare it to the diagonalization (eigenvalue decomposition) $A = UDU^T$ of a matrix.

Both decompositions provide a method of understanding a matrix through a diagonal matrix. Eigenvalue decomposition of a matrix is much more powerful and it gives a basis to which the matrix acts like scalar multiplication. On the other hand, singular value decomposition gives a basis of \mathbb{R}^n and a basis of \mathbb{R}^m , where the action of the matrix on the first basis can be understood through a scalar multiplication of the second basis.

Proof. Suppose that A has rank $r \leq \min\{m, n\}$. Note that $A^T \cdot A$ also has rank r . Let $\sigma_1^2, \dots, \sigma_r^2$ be the non-zero eigenvalues of $A^T \cdot A$, and let v_i be orthonormal eigenvectors (where v_i is the eigenvector corresponding to σ_i for $1 \leq i \leq r$). Complete the collection v_1, v_2, \dots, v_r into an orthonormal basis by adding vectors v_{r+1}, \dots, v_n .

For $1 \leq i \leq r$, let $u_i = \frac{Av_i}{\sigma_i}$. Note that for distinct i, j we have,

$$u_i^T \cdot u_j = \frac{1}{\sigma_i \sigma_j} v_j^T A^T A v_i = \frac{1}{\sigma_i \sigma_j} v_j^T \cdot (\sigma_i v_i) = 0,$$

since v_j and v_i are orthogonal. Therefore, u_1, \dots, u_r is an orthogonal set of vectors. Complete this collection as a set of orthonormal basis by adding vectors u_{r+1}, \dots, u_m and let $\sigma_i = 0$ for all $r+1 \leq i \leq m$. Let $U = [u_1 \ u_2 \ \dots \ u_m]$ and $V = [v_1 \ v_2 \ \dots \ v_n]$ and note that

$$U^T A V = U^T \cdot [\sigma_1 u_1 \ \sigma_2 u_2 \ \dots \ \sigma_m u_m] = [\sigma_1 e_1 \ \sigma_2 e_2 \ \dots \ \sigma_m e_m] = \Sigma,$$

is a diagonal matrix. Hence

$$A = U \Sigma V^T,$$

and we found our matrices U and V as claimed. \square

Definition 6. The (positive) square roots of the eigenvalues of $A^T A$ is called the singular values of A . It is customary to denote the singular values in decreasing order of magnitude $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_m$. Note that these are the diagonal entries of the matrix Σ_{mn} , in the decomposition given above.

Exercise 7. Prove that the singular values can also be obtained by considering the matrix AA^T .

Recall the example given in the introduction where we consider a matrix whose rows are indexed by stocks and columns by dates. Is it more sensible to consider an eigenvalue decomposition or a singular value decomposition?

We mostly will use matrices with $m \leq n$. In this case, we can further simplify the decomposition as follows.

Corollary 8. (Reduced SVD) Let $m \leq n$. Then every $m \times n$ matrix A_{mn} can be expressed as

$$A_{mn} = U_{mm} \Sigma_{mm} V_{nm}^T$$

for a $m \times m$ orthogonal matrix U_{mm} , a diagonal $m \times m$ matrix Σ_{mm} , and a $n \times m$ matrix V_{nm} whose columns are orthogonal.

Proof. If $m \leq n$, then in the diagonal matrix Σ_{mn} , for $j > m$, the j -th column of Σ_{mn} is a all-zero vector. Hence even if we remove the $(m + 1)$ -th column to the n -th column of Σ_{mn} , and remove the $(m + 1)$ -th column to the n -th column of V_{nm} , the outcome of

$$\Sigma_{mn} V_{nm}^T$$

does not change. Hence we obtain a decomposition as claimed. \square

Exercise 9. Are the matrices U and V uniquely determined?

Example 10. Consider the matrix

$$A = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix}.$$

We see that

$$AA^T = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & -2 \end{bmatrix} \cdot \begin{bmatrix} 3 & 2 \\ 2 & 3 \\ 2 & -2 \end{bmatrix} = \begin{bmatrix} 13 & 12 & 2 \\ 12 & 13 & -2 \\ 2 & -2 & 8 \end{bmatrix}.$$

The eigenvalues of this matrix are 0, 9, 25, and the singular values are 5, 3. We can also use $A^T A$, which is a 2×2 matrix, to obtain the singular values.

To obtain the singular value decomposition, we need to find the eigenvectors of AA^T . To find the eigenvector corresponding to the eigenvalue 25, note that

$$AA^T - 25I = \begin{bmatrix} -12 & 12 & 2 \\ 12 & -12 & -2 \\ 2 & -2 & -13 \end{bmatrix}.$$

The matrix above can be row-reduced to give the matrix

$$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

and therefore $(1, 1, 0)$ is an eigenvector corresponding to 25. By normalizing this vector, we can take $v_1 = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)$ as the eigenvector. Similarly, we can also find a unit eigenvector $v_2 = (\frac{1}{\sqrt{18}}, -\frac{1}{\sqrt{18}}, \frac{4}{\sqrt{18}})$ of 9. (There is no need to find the eigenvector corresponding to 0 when finding the simplified form of SVD).

In the proof of the singular value decomposition, the matrix U consisted of the vectors $u_i = \frac{Av_i}{\sigma_i}$. Therefore,

$$u_1 = \begin{bmatrix} 1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix} \quad u_2 = \begin{bmatrix} 1/\sqrt{2} \\ -1/\sqrt{2} \end{bmatrix}.$$

Hence,

$$U = [u_1 \ u_2] \quad \text{and} \quad V = [v_1 \ v_2],$$

and

$$A = U\Sigma V^T = \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ 1/\sqrt{2} & -1/\sqrt{2} \end{bmatrix} \cdot \begin{bmatrix} 5 & 0 \\ 0 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1/\sqrt{2} & 1/\sqrt{2} & 0 \\ 1/\sqrt{18} & -1/\sqrt{18} & 4/\sqrt{18} \end{bmatrix}.$$

4. THE PERRON-FROBENIUS THEOREM

Theorem 11. (*Perron-Frobenius*) *If A is an $n \times n$ matrix whose entries are all positive (greater than zero), then*

1. *There is a positive real eigenvalue λ_0 such that all other eigenvalues satisfy $|\lambda| < \lambda_0$,*
2. *there is a positive eigenvector corresponding to λ_0 , and*
3. *λ_0 is an eigenvalue of multiplicity 1.*

Proof. We only give the sketch of the proof for symmetric matrices.

Note that symmetric matrices have real eigenvalues. Let λ_+ be the largest eigenvalue and λ_- be the smallest eigenvalue. If $\lambda_+ > |\lambda_-|$, then Property 1 follows. Suppose not (thus $|\lambda_-| \geq \lambda_+$). Then by Theorem 2 above, it follows that the eigenvector v of λ_- maximizes the quantity

$$\frac{\|Av\|}{\|v\|}$$

over all non-zero vectors. The vector v has real entries. Without loss of generality assume that $v \cdot (1, 1, \dots, 1) \geq 0$. Since A has positive entries, whenever we switch the sign of a negative entry of v , we obtain a vector of the same norm as v having larger value of $\|Av\|$. This contradicts the fact that $|\lambda_-| \geq \lambda_+$. Therefore, we must have $\lambda_+ > |\lambda_-|$.

A similar argument shows that all entries of an eigenvector w corresponding to λ_+ needs to be non-negative. Moreover, since A has positive entries, the only way w can have a zero entry is if w is the all-zero vector. Hence we have Property 2.

Finally, to see Property 3, suppose that there are two linearly independent unit eigenvectors w_1 and w_2 both corresponding to λ_+ . Since $w_1 \neq w_2$ and w_1, w_2 are both unit vectors, we see that $w_1 - w_2$ must have both positive and negative entries. Moreover, $w_1 - w_2$ is also an eigenvector corresponding to λ_+ . This is impossible by the argument above saying that all entries of an eigenvector corresponding to λ_+ must have the same sign. \square

Perron-Frobenius theorem holds for a more general class of matrices (see, e.g., Wikipedia).

MIT OpenCourseWare
<http://ocw.mit.edu>

18.S096 Topics in Mathematics with Applications in Finance
Fall 2013

For information about citing these materials or our Terms of Use, visit: <http://ocw.mit.edu/terms>.