MATH-UA 140 - Linear Algebra

Lecture 22 : Diagonalizing a matrix

We have seen in the previous lecture that the eigenvalues of a matrix give us a "dynamical" understanding of a matrix, i.e the effect of the matrix when it is applied to the same vector multiple times. In this lecture we formalize this idea by learning how to diagonalize a matrix, i.e how to find a factorization for the matrix which shows its eigenvalues explicitly.

I] Diagonalization

1) Diagonalizing a matrix

Suppose A is an $n \times n$ matrix with n linearly independent eigenvectors $\vec{x}_1, \dots, \vec{x}_n$, with associated eigenvalues $\lambda_1, \dots, \lambda_n$ (which may not be distinct. Consider the eigenvector matrix S whose columns are the eigenvectors $\vec{x}_1, \dots, \vec{x}_n$. Then,

$$S^{-1} A S = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \Lambda$$

Λ, the diagonal matrix with the eigenvalues on the diagonal, is called the eigenvalue matrix.

$S^{-1}AS = \Lambda$ is called the diagonalization of A.

Let us see why this is true:

* If we multiply A times S, the first column of AS is $A\vec{x}_1$, the second column is $A\vec{x}_2$, etc.
Since $\vec{x}_1, \vec{x}_2, \dots, \vec{x}_n$ are eigenvectors, $A\vec{x}_1 = \lambda_1 \vec{x}_1$, $A\vec{x}_2 = \lambda_2 \vec{x}_2$, etc., so that

$$AS = \begin{bmatrix} \\ \lambda_1 \vec{x}_1 & \lambda_2 \vec{x}_2 & \cdots & \lambda_n \vec{x}_n \\ \\ \end{bmatrix}$$

* On the other hand, if we evaluate $S\Lambda$ we find

$$\begin{bmatrix} \\ \vec{x}_1 & \cdots & \vec{x}_n \\ \\ \end{bmatrix} \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \ddots & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{bmatrix} = \begin{bmatrix} \\ \lambda_1 \vec{x}_1 & \lambda_2 \vec{x}_2 & \cdots & \lambda_n \vec{x}_n \\ \\ \end{bmatrix}$$

Hence, $AS = S\Lambda$

* S is an $n \times n$ matrix with n linearly independent columns, so it is invertible. Multiplying the equality above by S^{-1}, we find the desired equality:

$$S^{-1}AS = \Lambda$$

Example: In the last lecture, we found that
$A = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}$ had two eigenvalues $\lambda_1 = -2$, and $\lambda_2 = 5$, with associated eigenvectors $\begin{bmatrix} -\frac{4}{3} \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$

The two eigenvectors are linearly independent, and we can write:

$$\begin{bmatrix} -\frac{4}{3} & 1 \\ 1 & 1 \end{bmatrix}^{-1} \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -\frac{4}{3} & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}$$

$$\Leftrightarrow \begin{bmatrix} -\frac{3}{7} & \frac{3}{7} \\ \frac{3}{7} & \frac{4}{7} \end{bmatrix} \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix} \begin{bmatrix} -\frac{4}{3} & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} -2 & 0 \\ 0 & 5 \end{bmatrix}$$

2) Important remarks

* The eigenvectors in S come in the same order as the eigenvalues in Λ. If you exchange the order in S, you have to exchange the order in Λ as well.

* Imagine you have a factorization $S^{-1}AS = \Lambda$, with Λ a diagonal matrix. Then S must be a matrix of eigenvectors. Indeed, the equality implies $AS = S\Lambda$.
The i^{th} column of AS is $A\vec{x_i}$, and the i^{th} column of $S\Lambda$ is $\lambda_i \vec{x_i}$, where λ_i is the i^{th} diagonal entry of Λ. That means that $\vec{x_i}$ is an eigenvector of A with eigenvalue λ_i.

* Imagine A is an $n \times n$ matrix with <u>n distinct eigenvalues</u> $\lambda_1, \dots, \lambda_n$. Then the eigenvectors $\vec{x_1}, \dots, \vec{x_n}$ corresponding to the distinct eigenvalues are <u>linearly independent</u>.

Here is the proof, first for a 2×2 matrix and 2 eigenvectors $\vec{x_1}$ and $\vec{x_2}$:
Let us imagine there are scalars c_1 and c_2 such that $c_1\vec{x_1} + c_2\vec{x_2} = 0$

If we apply A to this equality, we find:

$c_1 A\vec{x_1} + c_2 A\vec{x_2} = \vec{0} \iff \lambda_1 c_1 \vec{x_1} + \lambda_2 c_2 \vec{x_2} = \vec{0}$

Now, multiplying $c_1\vec{x_1} + c_2\vec{x_2} = \vec{0}$ by λ_2, we also have

$\lambda_2 c_1 \vec{x_1} + \lambda_2 c_2 \vec{x_2} = \vec{0}$

Subtracting the two equalities, we get: $(\lambda_1 - \lambda_2) c_1 \vec{x_1} = \vec{0}$

Since λ_1 and λ_2 are distinct, $c_1 = 0$, and then $c_2 = 0$

So $\vec{x_1}$ and $\vec{x_2}$ are linearly independent.

The proof generalizes to n eigenvectors easily

Suppose $c_1\vec{x_1} + c_2\vec{x_2} + \dots + c_n\vec{x_n} = \vec{0}$

Then applying A, $c_1\lambda_1\vec{x_1} + c_2\lambda_2\vec{x_2} + \dots + c_n\lambda_n\vec{x_n} = \vec{0}$

And multiplying the first equality by λ_n,

$c_1\lambda_n\vec{x_1} + c_2\lambda_n\vec{x_2} + \dots + c_n\lambda_n\vec{x_n} = \vec{0}$

Subtracting, $c_1(\lambda_1 - \lambda_n)\vec{x_1} + c_2(\lambda_2 - \lambda_n)\vec{x_2} + \dots + c_{n-1}(\lambda_{n-1} - \lambda_n)\vec{x_{n-1}} = \vec{0}$

We repeat the steps: multiply this equality $\nearrow$ by A, $\searrow$ by λ_{n-1}, and

subtract: $c_1(\lambda_1 - \lambda_n)(\lambda_1 - \lambda_{n-1})\vec{x_1} + c_2(\lambda_2 - \lambda_n)(\lambda_2 - \lambda_{n-1})\vec{x_2} + \dots + c_{n-2}(\lambda_{n-2} - \lambda_n)$
$(\lambda_{n-2} - \lambda_{n-1})\vec{x_{n-2}} = \vec{0}$

and we can continue the process until we find:

$(\lambda_1 - \lambda_n)(\lambda_1 - \lambda_{n-1}) \cdots (\lambda_1 - \lambda_2) c_1 \vec{x_1} = 0 \Rightarrow c_1 = 0$

and all the c_i must then be zero, going back one equality at a time.

<u>Bottom line</u>: Any matrix with n distinct eigenvalues can be diagonalized.

II) Mistakes to avoid with eigenvalues and diagonalization

1) Nondiagonalizable matrices

Some $n \times n$ matrices do not have n independent eigenvectors. These matrices cannot be diagonalized.
From what we have learned, this can only happen when a matrix does not have n distincts eigenvalues.

Example: $A = \begin{bmatrix} 3 & -1 \\ 1 & 1 \end{bmatrix}$ $\quad \det(A - \lambda I) = \begin{vmatrix} 3-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix}$

$$= (3-\lambda)(1-\lambda) + 1$$
$$= 3 - 4\lambda + \lambda^2 + 1 = (\lambda - 2)^2$$

The only eigenvalue of the matrix A is $\lambda = 2$. The eigenvectors are in the nullspace of $\begin{bmatrix} 1 & -1 \\ 1 & -1 \end{bmatrix}$. This is a one-dimensional subspace, with basis $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$. A does not have a second linearly independent eigenvector, so A is not diagonalizable.

The numbers of repetition of a given eigenvalue among the n roots of $\det(A - \lambda I) = 0$ is called the algebraic multiplicity of that eigenvalue. In the example above, the eigenvalue $\lambda = 2$ has algebraic multiplicity 2.
The number of linearly independent eigenvectors corresponding to a given eigenvalue is called the geometric multiplicity.

The geometric multiplicity is always less than or equal to the algebraic multiplicity. When it is strictly less than the algebraic multiplicity, as in our example, the matrix is not diagonalizable.

2) Eigenvalues of AB and $A+B$

Imagine we have two square matrices A and B, and that A has λ as an eigenvalue, and B has β as an eigenvalue.

In general, $\lambda\beta$ is NOT an eigenvalue of AB and $\lambda+\beta$ is NOT an eigenvalue of $A+B$.

It is tempting to write $AB\vec{x} = A(\beta\vec{x}) = \beta A\vec{x} = \lambda\beta\vec{x}$

and $(A+B)\vec{x} = A\vec{x} + B\vec{x} = (\lambda+\beta)\vec{x}$

however, those equalities do not hold in general, because A and B may not share the same eigenvector.

The equalities above are only true when A and B share an eigenvector $\vec{x}$.

III Matrix powers

1) Fast computation of powers of a matrix

Eigenvalues and diagonalization are very good mathematical tools to understand the cumulative effect of multiplying a matrix with a vector. To see this, assume a matrix A can be diagonalized: $A = S\Lambda S^{-1}$

Then,
$$A^2 = (S\Lambda S^{-1})(S\Lambda S^{-1}) = S\Lambda^2 S^{-1}$$
$$A^3 = AA^2 = (S\Lambda S^{-1})(S\Lambda^2 S^{-1}) = S\Lambda^3 S^{-1}$$
$$\vdots$$
$$A^k = AA^{k-1} = (S\Lambda S^{-1})(S\Lambda^{k-1} S^{-1}) = S\Lambda^k S^{-1}$$

The point here is to see that Λ is a diagonal matrix, so Λ^k is easily calculated: it is the diagonal matrix whose entries are the eigenvalues of Λ to the k^{th} power.

Example: $A = \begin{bmatrix} \frac{3}{10} & \frac{4}{10} \\ \frac{7}{10} & \frac{6}{10} \end{bmatrix}$

$$\begin{vmatrix} \frac{3}{10}-\lambda & \frac{4}{10} \\ \frac{7}{10} & \frac{6}{10}-\lambda \end{vmatrix} = \frac{1}{100}\begin{vmatrix} 3-10\lambda & 4 \\ 7 & 6-10\lambda \end{vmatrix}$$
$$= \frac{1}{100}\left[(3-10\lambda)(6-10\lambda) - 28\right]$$
$$= \frac{1}{10}\left[10\lambda^2 - 9\lambda - 1\right] = 0$$
$$\Rightarrow \lambda_1 = -\frac{1}{10} \qquad \lambda_2 = 1$$

The eigenvectors are obtained in the usual way:

For λ_1: $\begin{bmatrix} \frac{4}{10} & \frac{4}{10} \\ \frac{7}{10} & \frac{7}{10} \end{bmatrix}$ has $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ as a basis for its nullspace

$\begin{bmatrix} -\frac{7}{10} & \frac{4}{10} \\ \frac{7}{10} & -\frac{4}{10} \end{bmatrix}$ has $\begin{bmatrix} 1 \\ \frac{7}{4} \end{bmatrix}$ as a basis for its nullspace

$$S = \begin{bmatrix} -1 & 1 \\ 1 & \frac{7}{4} \end{bmatrix} \longrightarrow S^{-1} = -\frac{4}{11} \begin{bmatrix} \frac{7}{4} & -1 \\ -1 & -1 \end{bmatrix} = \begin{bmatrix} -\frac{7}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} \end{bmatrix}$$

and the diagonalization of A is:

$$A = \begin{bmatrix} -1 & 1 \\ 1 & \frac{7}{4} \end{bmatrix} \begin{bmatrix} -0.1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{7}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} \end{bmatrix}$$

Any power of A can now be easily computed. For example,

$$A^4 = \begin{bmatrix} -1 & 1 \\ 1 & \frac{7}{4} \end{bmatrix} \begin{bmatrix} -0.0001 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{7}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} \end{bmatrix}$$

More generally,

$$A^k = \begin{bmatrix} -1 & 1 \\ 1 & \frac{7}{4} \end{bmatrix} \begin{bmatrix} -0.1^k & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{7}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} \end{bmatrix}$$

In the limit $k \to \infty$:

$$A^\infty = \begin{bmatrix} -1 & 1 \\ 1 & \frac{7}{4} \end{bmatrix} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} -\frac{7}{11} & \frac{4}{11} \\ \frac{4}{11} & \frac{4}{11} \end{bmatrix}$$

In this way, we also see that if A is diagonalizable, $A^k \to 0$ if all its eigenvalues λ are such that $|\lambda|<1$

zero matrix

2) Fast computation of sequences of vectors

Let us imagine there is a sequence of vectors $\vec{u}_k$ given by $\vec{u}_k = A\vec{u}_{k-1}$ starting at some $\vec{u}_0$. $\vec{u}_k$ can be computed in terms of $\vec{u}_0$ directly:
$\vec{u}_k = A\,\vec{u}_{k-1} = A^2\,\vec{u}_{k-2} = \cdots = A^k\,\vec{u}_0$

If A is diagonalizable, then $\vec{u}_k = S\Lambda^k S^{-1}\vec{u}_0$. This is already quite nice. However, there is an even more intuitive way to visualize and compute this expression.

- First, write $\vec{u}_0$ in the eigenvector basis:
  $\vec{u}_0 = c_1\vec{x}_1 + c_2\vec{x}_2 + \cdots + c_n\vec{x}_n$
  This is $\vec{u}_0 = S\vec{c} \iff \vec{c} = S^{-1}\vec{u}_0$
- Multiply each c_i by λ_i^k.
  This corresponds to $\Lambda^k\vec{c} = \Lambda^k S^{-1}\vec{u}_0$
- Construct the linear combination of the $\vec{x}_i$ with $\lambda_i^k c_i$ as the multiplying scalars
  This is $S\Lambda^k\vec{c} = S\Lambda^k S^{-1}\vec{u}_0 = \vec{u}_k$

We conclude that $\vec{u}_k = A\vec{u}_{k-1} = A^k\vec{u}_0 = c_1\lambda_1^k\vec{x}_1 + c_2\lambda_2^k\vec{x}_2 + \cdots + c_n\lambda_n^k\vec{x}_n$

Example Let $\vec{u}_0 = (1,0)$. Compute $A^k\vec{u}_0$ for $A = \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix}$

which has $\lambda = -1$ and $\lambda = 3$ as eigenvalues, and $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$ and $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$ as eigenvectors.

We have $\vec{u}_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$

$= c_1 \vec{x}_1 + c_2 \vec{x}_2$ with $c_1 = -\frac{1}{2}$, $c_2 = \frac{1}{2}$

$$\vec{u}_k = -\frac{1}{2} (-1)^k \begin{bmatrix} -1 \\ 1 \end{bmatrix} + \frac{1}{2} 3^k \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$