Diagonalizable Matrices

Richard Ketchersid

Grand Canyon University

Diagonal Matrices

Diagonal matrices have many nice properties that make them behave almost as scalars. Some of the properties are listed here, then generalized to diagonalizable matrices.

Commutativity

If A and B are diagonal nxn matrices, then AB is the diagonal $n \times n$ matrix given by:

$$\begin{bmatrix} A_{11} & 0 & \cdots & 0 \\ 0 & A_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn} \end{bmatrix} \begin{bmatrix} B_{11} & 0 & \cdots & 0 \\ 0 & B_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & B_{nn} \end{bmatrix} = \begin{bmatrix} A_{11}B_{11} & 0 & \cdots & 0 \\ 0 & A_{22}B_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn}B_{nn} \end{bmatrix}$$

From this and commutativity of multiplication in the scalar field is is clear that $AB = BA$. Hence diagonal matrices commute.

Inverse

If A is an $n \times n$ diagonal matrix with no 0's on the diagonal, then A^{-1} is given by

$$\begin{bmatrix} A_{11}^{-1} & 0 & \cdots & 0 \\ 0 & A_{22}^{-1} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn}^{-1} \end{bmatrix} = \begin{bmatrix} \frac{1}{A_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{A_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{A_{nn}} \end{bmatrix}$$

If there is a 0 on the diagonal, then A is not invertible.

Powers

If A is an $n \times n$ diagonal matrix and k any integer, then A^k is given by

$$\begin{bmatrix} A_{11}^k & 0 & \cdots & 0 \\ 0 & A_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_{nn}^k \end{bmatrix}$$

For $k < 0$, this requires that none of the diagonal entries of A are 0. This generalizes to k rational, real, or even complex.

Polynomials

If A is an $n \times n$ diagonal matrix and $p(x) = \sum_{i=0}^{k} c_i x^i$ be a polynomial, then $p(A) = \sum_{i=0}^{k} c_i x^i$ is given by:

$$\begin{bmatrix} \sum_{i=0}^{k} c_i A_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{k} c_i A_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{k} c_i A_{nn}^i \end{bmatrix}$$

So it is a very simple matter to generalize polynomials to act on diagonal matrices. Of course polynomial can generalize to arbitrary square matrices, but the result is not so simple.

General Matrix

If A is an $m \times n$ matrix, then

$$||A|| = \sup\{||A\boldsymbol{x}|| \mid ||\boldsymbol{x}|| = 1\}$$

Here $||\boldsymbol{x}|| = ||\boldsymbol{x}||_2 = \sqrt{\langle \boldsymbol{x}, \boldsymbol{x} \rangle}$.
An equivalent definition is

$$||A|| = \sup\left\{\frac{||A\boldsymbol{x}||}{||\boldsymbol{x}||} \,\middle|\, \boldsymbol{x} \in \mathbb{R}^n\right\}$$

So $||A\boldsymbol{x}|| \leq ||A||\,||\boldsymbol{x}||$ and $||A||$ is the least γ such that $||A\boldsymbol{x}|| \leq \gamma\,||\boldsymbol{x}||$ for all $\boldsymbol{x}$.

SVD will provide a nice and intuitive geometrical meaning to $||A||$.

Diagonal Matrix

If A is an $n \times n$ diagonal matrix, then $A\boldsymbol{x} = (A_{11}x_1, A_{12}x_2, \ldots, A_{nn}x_n)$ and clearly $||\boldsymbol{A}x|| \leq \max\{|A_{ii}| \mid 1 \leq i \leq n\}$ for $||\boldsymbol{x}|| = 1$.

It is then easy to see in this case that that

$$||A|| = \max\{|A_{ii}| \mid 1 \leq i \leq n\}.$$

Power Series

If A is an $n \times n$ diagonal matrix and $p(x) = \sum_{i=0}^{\infty} c_i x^i$ be a power series with radius of convergence R and $||A|| < R$, then $p(A) = \sum_{i=0}^{\infty} c_i x^i$ is given by:

$$\begin{bmatrix} \sum_{i=0}^{\infty} c_i A_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} c_i A_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} c_i A_{nn}^i \end{bmatrix}$$

So again it is a simple matter to generalize power series to act on diagonal matrices after appropriate generalization of norm. Of course power series can generalize to arbitrary square matrices, but the result is even harder to understand than the case for polynomials.

Power Series/Example 1

Recall the most basic example of power series:

$$\frac{1}{1-x} = \sum_{i=0}^{\infty} x^i \text{ for } |x| < 1$$

If A is an $n \times n$ diagonal matrix and $||A|| < 1$, then $(1-A)^{-1} = \sum_{i=0}^{\infty} A^i$. From earlier slides we see that this boils down to

$$\begin{bmatrix} \frac{1}{1-A_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-A_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{1-A_{nn}} \end{bmatrix} = \begin{bmatrix} \sum_{i=0}^{\infty} A_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} A_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} A_{nn}^i \end{bmatrix}$$

Power Series/Example 2

Another basic example of power series is:

$$e^x = \exp(x) = \sum_{i=0}^{\infty} \frac{1}{i!} x^i \text{ for all } x$$

If A is an $n \times n$ is any diagonal matrix $e^A = \exp(A)$ is defined by:

$$\sum_{i=0}^{\infty} \frac{1}{i!} A^i = \begin{bmatrix} \sum_{i=0}^{\infty} \frac{1}{i!} A_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} \frac{1}{i!} A_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} \frac{1}{i!} A_{nn}^i \end{bmatrix} = \begin{bmatrix} e^{A_{11}} & 0 & \cdots & 0 \\ 0 & e^{A_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{A_{nn}} \end{bmatrix}$$

Functions of Matrices

It is quite simple to "lift" a large class of functions from the world of scalars to the world of diagonal matrices with the result being:

$$f(A)_{ii} = f(A_{ii}) \text{ and } f(A)_{ij} = 0 \text{ if } i \neq j$$

It would be quite difficult to say anything useful about e^A if A failed to have some nice structure, like being diagonal. In addition, multivariable functions also lift simply, it is very simple to make sense of $f(A, B) = A^2B - AB^2$ for A and B diagonal $n \times n$ matrices. In fact the result is the diagonal matrix whose ii^{th} entry is just $f(A_{ii}, B_{ii})$:

$$f(A, B)_{ii} = A_{ii}^2 B_{ii} - A_{ii} B_{ii}^2 = f(A_{ii}, B_{ii}).$$

Definition

A square matrix A is *diagonalizable* iff $A = SDS^{-1}$ for some diagonal matrix D and some invertible matrix S.

Letting $S = [\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots,\boldsymbol{v}_n]$, this says that we can take the basis $\mathcal{V} = \{\boldsymbol{v}_1\ \boldsymbol{v}_2\ \cdots,\boldsymbol{v}_n\}$ and by representing the action of A in the basis $\mathcal{V}$, the resulting matrix is diagonal. In particular, $[A]_{\mathcal{V},\mathcal{V}} = D$.

Eigenvalue/Eigenvector

D_{ii} is an *eigenvalue* of A and $\boldsymbol{v}_i$ is an *eigenvector* corresponding to D_{ii}. $\mathcal{V}$ is a basis of eigenvectors.

Powers/Inverse

If A is an $n \times n$ diagonalizable matrix with $A = SDS^{-1}$ and k any integer, then

$$A^k = (SDS^{-1})^k = (SDS^{-1})(SDS^{-1}) \cdots (SDS^{-1}) = SD^k S^{-1}$$

all the $S^{-1}S$ pairs cancelling. So

$$A^k = S \begin{bmatrix} D_{11}^k & 0 & \cdots & 0 \\ 0 & D_{22}^k & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_{nn}^k \end{bmatrix} S^{-1}$$

Thus it is much easier, quicker and numerically stable to compute powers of diagonalizable matrices.

Polynomials

If A is an $n \times n$ diagonalizable matrix with $A = SDS^{-1}$ and $p(x) = \sum_{i=0}^{n} c_i x^i$ be a polynomial, then

$$p(A) = \sum_{i=0}^{k} c_i A^i = \sum_{i=0}^{k} c_i S D^i S^{-1} = S \left(\sum_{i=0}^{k} c_i D^i \right) S^{-1} = S\, p(D)\, S^{-1}$$

so

$$p(A) = S \begin{bmatrix} \sum_{i=0}^{k} c_i D_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{k} c_i D_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{k} c_i D_{nn}^i \end{bmatrix} S^{-1}$$

Power Series

If A is an $n \times n$ diagonalizable matrix with $A = SDS^{-1}$ and $p(x) = \sum_{i=0}^{\infty} c_i x^i$ be a power series with radius of convergence R, then

$$p(A) = \sum_{i=0}^{\infty} c_i A^i = \sum_{i=0}^{\infty} c_i S D^i S^{-1} = S \left(\sum_{i=0}^{\infty} c_i D^i \right) S^{-1} = S\, p(D)\, S^{-1}$$

provided $||D|| < R$. So

$$p(A) = S \begin{bmatrix} \sum_{i=0}^{\infty} c_i D_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} c_i D_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} c_i D_{nn}^i \end{bmatrix} S^{-1}$$

Power Series/Example 1

If A is an $n \times n$ diagonalizable matrix with $A = SDS^{-1}$ and $||D|| < 1$, then

$$(I - A)^{-1} = (SS^{-1} - SDS^{-1})^{-1} = S(I - D)^{-1}S^{-1} = \sum_{i=0}^{\infty} A^n = S\left(\sum_{i=0}^{\infty} D^i\right) S^{-1}$$

. From earlier slides we see that this boils down to

$$S \begin{bmatrix} \frac{1}{1-D_{11}} & 0 & \cdots & 0 \\ 0 & \frac{1}{1-D_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{1}{1-D_{nn}} \end{bmatrix} S^{-1} = S \begin{bmatrix} \sum_{i=0}^{\infty} D_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} D_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} D_{nn}^i \end{bmatrix} S^{-1}$$

Power Series/Example 2

If A is an $n \times n$ is any diagonalizable matrix with $A = SDS^{-1}$, then $e^A = \exp(A)$ is defined by:

$$\sum_{i=0}^{\infty} \frac{1}{i!} A^n = \sum_{i=0}^{\infty} \frac{1}{i!} SD^n S^{-1} = S \left(\sum_{i=0}^{\infty} \frac{1}{i!} D^n \right) S^{-1}$$

So e^A is given by:

$$S \begin{bmatrix} \sum_{i=0}^{\infty} \frac{1}{i!} A_{11}^i & 0 & \cdots & 0 \\ 0 & \sum_{i=0}^{\infty} \frac{1}{i!} A_{22}^i & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sum_{i=0}^{\infty} \frac{1}{i!} A_{nn}^i \end{bmatrix} S^{-1} = S \begin{bmatrix} e^{A_{11}} & 0 & \cdots & 0 \\ 0 & e^{A_{22}} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{A_{nn}} \end{bmatrix} S^{-1}$$

Functions of Matrices

It is quite simple to "lift" a large class of functions from the world of scalars to the world of diagonalizable matrices $A = SDS^{-1}$ with the result being:

$$f(A) = S \begin{bmatrix} f(D_{11}) & 0 & \cdots & 0 \\ 0 & f(D_{22}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f(D_{nn}) \end{bmatrix} S^{-1}$$

Functions of Matrices

Call $A_1, A_2, \ldots, A_k$ *jointly diagonalizable* if there is S so that $A_i = SD_iS^{-1}$ where D_i is diagonalizable, for all $i = 1, \ldots, k$. Then it often works to define $f(A_1, \ldots, A_k)$ by:

$$S \begin{bmatrix} f((D_1)_{11}, \ldots, (D_k)_{11}) & 0 & \cdots & 0 \\ 0 & f((D_1)_{22}, \ldots, (D_k)_{22}) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & f((D_1)_{nn}, \ldots, (D_k)_{nn}) \end{bmatrix} S^{-1}$$