Lecture IV
Eigenvalues, Eigenvectors, and Norms

I. Eigenvalues and Eigenvectors

Just to make sure that you haven't picked up any bad habits, the determinant of any *n**n* matrix can be derived by expanding down any column or across any row of the matrix:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = (-1)^{1+1} a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + (-1)^{2+1} a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{23} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} + (-1)^{3+1} a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix}$$

$$+ (-1)^{4+1} a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{23} \\ a_{32} & a_{33} & a_{34} \end{vmatrix}$$

where

$$\begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} = a_{22} \begin{vmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{vmatrix} - a_{32} \begin{vmatrix} a_{23} & a_{24} \\ a_{42} & a_{43} \end{vmatrix} + a_{42} \begin{vmatrix} a_{23} & a_{24} \\ a_{33} & a_{34} \end{vmatrix}.$$

The eigenvalues of a matrix *A* are then defined as the solutions to the equation

$$|A - \lambda I| = 0$$

where *I* is the identity matrix. One example of the use of eigenvalues is from differential equations:

$$\begin{aligned} \frac{dv}{dt} &= 4v - 5w \quad v = 8 \text{ at } t = 0 \\ \frac{dw}{dt} &= 2v - 3w \quad w = 8 \text{ at } t = 0 \end{aligned}.$$

Another way to write this problem is

$$u(t) = \begin{pmatrix} v(t) \\ w(t) \end{pmatrix}, u_0 = \begin{pmatrix} 8 \\ 5 \end{pmatrix}, A = \begin{pmatrix} 4 & -5 \\ 2 & -2 \end{pmatrix}.$$

Thus,

$$\frac{du}{dt} = Au(t),\ u(0) = u_0.$$

We are looking for solutions of the form

$$\begin{aligned} u(t) &= e^{\lambda t} y \\ w(t) &= e^{\lambda t} z \end{aligned}.$$

Substituting this solution into the system of differential equations, we have

$$\begin{pmatrix} \lambda e^{\lambda t} y \\ \lambda e^{\lambda t} z \end{pmatrix} = \begin{pmatrix} 4e^{\lambda t} y - 5e^{\lambda t} z \\ 2e^{\lambda t} y - 3e^{\lambda t} z \end{pmatrix}.$$

next, we divide through by $e^{\lambda t}$ yielding

$$\begin{pmatrix} \lambda y \\ \lambda z \end{pmatrix} = \begin{pmatrix} 4y - 5z \\ 2y - 3z \end{pmatrix}$$

or

$$\begin{aligned} \lambda \begin{pmatrix} y \\ z \end{pmatrix} &= \begin{pmatrix} 4 & -5 \\ 2 & -3 \end{pmatrix} \begin{pmatrix} y \\ z \end{pmatrix} \\ \lambda x &= Ax \end{aligned}$$

which is the definition of the eigenvalues, λ, and the eigenvectors, x,

$$\begin{aligned} Ax - \lambda x &= 0 \\ (A - \lambda I)x &= 0 \end{aligned}$$

where the eigenvalues of x lie in the null space of the matrix

$$A - \lambda I = \begin{pmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{pmatrix}.$$

More intuitively, we know that the matrix

$$\begin{pmatrix} 4-\lambda & -5 \\ 2 & -3-\lambda \end{pmatrix}$$

is singular by definition of the eigenvalue.

Given the existence of eigenvalues and eigenvectors, it is possible to derive the diagonal form of the matrix

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

where S is a matrix whose columns are the eigenvectors

$$S=\left(x_1\cdots x_n\right)$$

$$\begin{aligned}AS&=\left(\lambda_1 x_1\cdots\lambda_n x_n\right)\\ &=\left(x_1\cdots x_n\right)\begin{pmatrix}\lambda_1 & \cdots & 0\\ \vdots & \ddots & \vdots\\ 0 & \cdots & \lambda_n\end{pmatrix}.\end{aligned}$$

Therefore, $S^{-1}S= I$, $AS = S\Lambda$, $S^{-1}AS=\Lambda$, and $A = S\Lambda S^{-1}$.

Notationally, in Gill, Murray and Wright the eigenvalues of A are written as $\left\{\lambda_i\left[A\right]\right\}$. $P_i(A)$ is the product of the eigenvalues, and $\Pi(AB) =\Pi(A)\,\Pi(B)$. If A is symmetric (i) all the eigenvalues of A are real numbers, and (ii) the matrix has n distinct egenvalues. If A is nonsingular then all its eigenvalues are nonzero and the eigenvalues of A^{-1} are the reciprocals of the eigenvalues of A. The maximum and minimum eigenvalues satisfy

$$\begin{aligned}\lambda_{\max}\left[A\right]&=\max_{x\neq 0}\frac{x'Ax}{x'x}\\ \lambda_{\min}\left[A\right]&=\min_{x\neq 0}\frac{x'Ax}{x'x}\end{aligned}.$$

II. Norms

A vector norm which will be denoted $\|x\|$, must satisfy three properties:

1.) $\|x\|\geq 0$ for all x, and $\|x\|=0$ if and only if x is the zero vector,

2.) For any real number δ, $\|\delta x\|=|\delta|\,\|x\|$, and

3.) For any two vectors x and y, $\|x+y\|\leq\|x\|+\|y\|$.

Our basic definition of the p-norm is

$$\|x\|_p=\left(\sum_{i=1}^n |x_i|^p\right)^{1/p}.$$

From our discussion of the optimality conditions, we derived

$$f(x) - f(x^*) = \nabla_x f(x)\,dx + \tfrac{1}{2}\,dx\,\nabla_{xx}^2 f(x)\,dx' \le 0$$

which implies for optimality that

$$\left\| \nabla_x f(x) \right\|_p \to 0.$$

The matrix norm is analogously defined with the additional property that
4.) $\|AB\| \le \|A\|\|B\|$.

Another value of the norm is a discussion of the stability of the matrix. For example, what are the properties of the solution of $Ax=b$ if we perturb the x vector?

$$\begin{aligned} A(x+\delta x) &= (b+\delta b) \\ x+\delta x &= A^{-1}b + A^{-1}\delta b \end{aligned}.$$

Noting that $x=A^{-1}b$ we subtract x and take the norm of each side:

$$\|\delta x\| = \|A^{-1}\|\|\delta b\|.$$

Given that $\|b\| \le \|A\|\|x\|$ by the original equation, we have

$$\frac{1}{\|x\|} \le \frac{\|A\|}{\|b\|}.$$

Multiplying both side by the inequality involving the perturbation, δ, we get

$$\frac{\|\delta x\|}{\|x\|} \le \|A\|\|A^{-1}\|\frac{\|\delta b\|}{\|b\|}.$$

This number $\|A\|\|A^{-1}\|$ is the condition number of A denoted cond(A). Since

$$1 = \|I\| \le \|A\|\|A^{-1}\|$$

cond(A) ≥ 1. If the condition number of A is large, the matrix is ill-conditions. If the condition number is small, the matrix is well conditioned.

III. Matrix Factorizations

The matrix factorization $A=LU$ involves reducing a matrix into a lower diagonal matrix L and an upper diagonal matrix U. The LU factorization is computed by a set of elementary row operations. Taking the matrix

$$A = \begin{pmatrix} 2 & 1 & 0 \\ 6 & 2 & 6 \\ -4 & -3 & 9 \end{pmatrix}$$

the first elementary matrix is

$$E_1 = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & 0 & 1 \end{pmatrix}.$$

The resulting matrix is

$$A_1 = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 6 \\ 0 & -1 & 9 \end{pmatrix},$$

the second factorization is then

$$E_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix}$$

which yields the upper triangular matrix

$$A_2 = U = \begin{pmatrix} 2 & 1 & 0 \\ 0 & -1 & 6 \\ 0 & 0 & 3 \end{pmatrix}.$$

The lower triangular matrix is then derived as

$$L = E_1^{-1} E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -3 & 1 & 0 \\ 2 & -1 & 1 \end{pmatrix}.$$

The usefulness of the decomposition is twofold. First, the LU factorization is a quick way to solve large systems of equations. First, note that the solution of the system of linear equations $Ax=b$ for $b=(1,5,3)'$ is $x=(-5/2,4,1)'$. If we start with the augmented system of equations:

$$\left(A|b\right)=\left(\begin{array}{ccc|c}2 & 1 & 0 & 1\\ 6 & 2 & 6 & 5\\ -4 & -3 & 9 & 3\end{array}\right)$$

operating using the same sequence as the row operations above yields

$$\left(A|b\right)=\left(\begin{array}{ccc|c}2 & 1 & 0 & 1\\ 0 & -1 & 6 & 2\\ 0 & 0 & 3 & 3\end{array}\right).$$

In this system, the solution of the original equations is accomplished by taking the last equation first:

$$\begin{aligned}3x_3 &= 3\\ x_3 &= 1\end{aligned}.$$

Given that $x_3 = 1$ the next equation up becomes

$$\begin{aligned}-x_2+6(1) &= 2\\ x_2 &= 4\end{aligned}.$$

The first equation then completes the solution

$$\begin{aligned}2x_1+(4) &= 1\\ x_1 &= -\frac{5}{2}\end{aligned}.$$

The second major use is in the cholesky decomposity. If a matrix is symmetric and positive definite, it can be decomposed so that $A=PP'$. Using the matrix

$$A=\begin{pmatrix}7 & 3 & 1\\ 3 & 4 & 0\\ 1 & 0 & 2\end{pmatrix}.$$

First, we need to determine whether the matrix is positive definite. The eigenvalues of the matrix are 2.5572, 1.4841, and 8.9587 so the matrix is positive definite. Following the procedure above, we derive the L matrix as

$$L=\begin{pmatrix}1.0000 & 0.0000 & 0.0000\\ 0.4286 & 1.0000 & 0.0000\\ 0.1429 & -0.1579 & 1.0000\end{pmatrix}$$

with the upper triangular matrix as

$$U = \begin{pmatrix} 7.0000 & 3.0000 & 1.0000 \\ 0.0000 & 2.7143 & -0.4286 \\ 0.0000 & 0.0000 & 1.7895 \end{pmatrix}.$$

Noting that the diagonal of the L matrix is 1 and the diagonal of the U matrix are greater than zero, suppose that we constructed the matrix

$$D^{-1} = \begin{pmatrix} 0.1429 & 0.0000 & 0.0000 \\ 0.0000 & 0.3684 & 0.0000 \\ 0.0000 & 0.0000 & 0.5588 \end{pmatrix}$$

which is one over the diagonal elements of the U matrix. Multiplying D^{-1} U yields L'.