Roundoff errors can be magnified by a poor choice of pivot when using Gaussian elimination.

Fact: Usually, the effect of roundoff is reduced by choosing, from the column in which the pivot is to come, an entry of largest absolute value to use as pivot.

This modification to Gaussian elimination gives the algorithm for Gaussian elimination with *partial pivoting*.

following, rounding off to 3 significant figures:

$$0.0001x + y = 1 \qquad (1)$$
$$x + y + 2 \qquad (2)$$

7. Eigenvalues and eigenvectors

Def'n: Let A be a square matrix. Then an *eigenvector* of A is a vector $\boldsymbol{v} \neq \mathbf{0}$ such that

$$A\boldsymbol{v} = \lambda\boldsymbol{v},$$

for some scalar λ.
The scalar λ is called the corresponding *eigenvalue*.

If $\boldsymbol{v}$ is an eigenvector of A, then so is $t\boldsymbol{v}$ for any scalar $t \neq 0$.

Example: Check with $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ and with $\boldsymbol{v} = (1\ \ 0\ -1)^T$, and $(1\ 2\ 1)^T$ and $(1\ \ -1\ \ 1)^T$.

Recall: If λ is an eigenvalue of A, with corresponding eigenvector $\boldsymbol{v}$, then $A\boldsymbol{v} = \lambda\boldsymbol{v} = \lambda I\boldsymbol{v}$, so $(A - \lambda I)\boldsymbol{v} = \mathbf{0}$. Hence $\boldsymbol{x} = \boldsymbol{v}$ is a non-trivial solution to the homogeneous system of equations $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$, and conversely, if there's a non-trivial sol'n then λ is an eigenvalue of A.
Thus:

λ is an eigenvalue of A
if and only if $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$ has a nontrivial sol'n
if and only if $A - \lambda I$ is singular
if and only if $\det(A - \lambda I) = 0$.

For $n \times n$ matrix A, $\det(A - \lambda I)$ is a polynomial of degree n in λ; called the *characteristic polynomial* of A.

The equation $\det(A - \lambda I) = 0$ is the *characteristic equation* of A.

Eigenvalues λ may be complex numbers, and the eigenvalues $\boldsymbol{v}$ may have complex components, even for real matrices A.

For $n = 2, 3$, solve the char. equ'n to get eigenvalues.
For $n \geq 4$ there are better numerical methods.

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

1. A and A^T have the same eigenvalues.

2. A is singular iff $\lambda = 0$ is an eigenvalue of A.

3. If λ is an eigenvalue of A, then λ^2 is an eigenvalue of A^2, and $1/\lambda$ is an eigenvalue of A^{-1} when A is non-singular.

4. If λ is an eigenvalue of A, then $\lambda - m$ is an eigenvalue of $A - mI$, for any scalar m.

Example

Earlier example of 3 masses joined by springs: had $\ddot{\boldsymbol{x}} = A\boldsymbol{x}$, where $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$. Eigenvalues were $-3, -1, -4$ with corresponding eigenvectors $\boldsymbol{v}_1 = (1\ 0\ -1)^T$, $\boldsymbol{v}_2 = (1\ 2\ 1)^T$, $\boldsymbol{v}_3 = (1\ -1\ 1)^T$.

Form matrix $P = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & -1 \\ -1 & 1 & 1 \end{pmatrix}$ whose cols are these eigenvectors.

Claim: $P^{-1}AP = \begin{pmatrix} -3 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -4 \end{pmatrix} = D$ say.

(D is diagonal with eigenvalues on the diagonal.)

No accident! And check that $AP = PD$.

8. Diagonalization

Def'n A square matrix A is *diagonalizable* if there is a non-singular matrix P such that $P^{-1}AP$ is a diagonal matrix.

Def'n Two matrices A and B are *similar* if there is a non-singular matrix P such that $B = P^{-1}AP$.

So A is diagonalizable
$\leftrightarrow$ A is similar to a diagonal matrix.

Theorem: Similar matrices have the same eigenvalues.
In fact: If $B = P^{-1}AP$ and $\boldsymbol{v}$ is an eigenvector of A corresponding to eigenvalue λ, then $P^{-1}\boldsymbol{v}$ is an eigenvector of B corresponding to eigenvalue λ.

independent eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$, corresponding to eigenvalues $\lambda_1, \ldots, \lambda_n$. Let

$$P = (\boldsymbol{v}_1 \ldots \boldsymbol{v}_n)$$

be the $n \times n$ matrix whose columns are the eigenvectors. Then

$$P^{-1}AP = \begin{pmatrix} \lambda_1 & 0 & \ldots & 0 \\ 0 & \lambda_2 & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \lambda_n \end{pmatrix},$$

the diagonal matrix with the eigenvalues down the main diagonal.

Let $A = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Easy to see the char. equ'n of both A and B is $(2-\lambda)(1-\lambda)^2 = 0$, so $\lambda = 2, 1, 1$.

Theorem: (Diagonalizability condition)
The $n \times n$ matrix A is diagonalizable if and only if A has n lin. indep. eigenvectors.

Theorem: Let A be $n \times n$ and let $\lambda_1, \ldots, \lambda_m$ be *distinct* eigenvalues of A, with corresponding eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$. Then $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m$ are linearly independent.

Thus eigenvectors corresponding to *different* eigenvalues are automatically lin. independent.

HENCE: If A is $n \times n$ with n distinct eigenvalues, then A is diagonalizable.

Example:
Say λ_1 and λ_2 (with $\lambda_1 \neq \lambda_2$) are eigenvalues of A, with corresponding e'vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$. So $A\boldsymbol{v}_1 = \lambda_1 \boldsymbol{v}_1$ and $A\boldsymbol{v}_2 = \lambda_2 \boldsymbol{v}_1$.

To check that $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are lin. indep., for what x_1, x_2 can we have

$$x_1\boldsymbol{v}_1 + x_2\boldsymbol{v}_2 = \mathbf{0} \qquad (1)$$

Find only sol'n is $x_1 = x_2 = 0$, so $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are lin. indep.

Applications of diagonalizability

1. Systems of differential equations

For a system of coupled differential equations which can be written in matrix form as

$$\dot{\boldsymbol{x}} = A\boldsymbol{x} \qquad (1)$$

(where $\boldsymbol{x} = (x_1, \ldots, x_n)^T$, $\dot{\boldsymbol{x}} = (\dot{x}_1, \ldots, \dot{x}_n)^T$),
if A can be diagonalized, say $P^{-1}AP = D$ with D diagonal, then make the substitution $\boldsymbol{x} = P\boldsymbol{y}$ in (1). This yields:

$$\dot{\boldsymbol{y}} = D\boldsymbol{y} \qquad (2)$$

which is easily solved.

2. Matrix powers

If A is diagonalizable, say $P^{-1}AP = D$ with D diagonal, then

$$A^n = PD^nP^{-1}.$$

This gives an easy way to calculate A^n.

$$\begin{aligned} \dot{x}_1 &= -3x_1 + x_2 \\ \dot{x}_2 &= x_1 - 2x_2 + x_3 \\ \dot{x}_3 &= x_2 - 3x_3. \end{aligned}$$

For $A = \begin{pmatrix} 1 & 6 \\ 1 & 0 \end{pmatrix}$, find a formula for A^n.

Difference Equations

Have a sequence of numbers $x_0, x_1, x_2, \ldots$ which satisfy a *difference equation*; e.g.

$$x_{n+2} = 3x_{n+1} - 2x_n \qquad \text{for } n \geq 0.$$

Simplest is 1st order: $x_{n+1} = ax_n$, (a constant).

Simultaneous (or coupled) first order:

Have two or more sequences, $x_0, x_1, x_2, \ldots,\ y_0, y_1, y_2, \ldots$ which satisfy

$$\begin{aligned} x_{n+1} &= ax_n + by_n \\ y_{n+1} &= cx_n + dy_n \end{aligned}$$

(here a, b, c, d are given constants). Write as a matrix equation: put $\boldsymbol{x}_n = \begin{pmatrix} x_n \\ y_n \end{pmatrix}$, $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Equations become: $\boldsymbol{x}_{n+1} = A\boldsymbol{x}_n$.

Then sol'n is: $\boldsymbol{x}_n = A^n \boldsymbol{x}_0$.

Find A^n by diagonalizing A.

One second order equation:

$x_{n+2} = ax_{n+1} + bx_n$ (where a, b are given constants).

Example: Find sequence $\{x_i\}_{i \geq 0}$, given $x_0 = 1$, $x_1 = 2$ and $x_{n+2} = x_{n+1} + 6x_n$ for $n = 0, 1, 2, \ldots$.

Note: When $n \times n$ matrix A is NOT diagonalizable, we find that A is similar to an $n \times n$ matrix with the eigenvalues (including repeats) on the main diagonal, and the rest 0 *except* for some 1's on the super-diagonal. E.g. like $\begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Called a *Jordan normal form* for A.

Def'n: An *orthogonal* matrix is a real square matrix Q such that the columns of Q are mutually orthogonal unit vectors
(i.e. $\boldsymbol{v}_i \cdot \boldsymbol{v}_j = o$ if $i \neq j$, and $\|\boldsymbol{v}_i\|_2 = 1$).

Equivalently, an orthogonal matrix is a real square matrix Q such that $Q^{-1} = Q^T$.

Recall: a matrix A is *symmetric* when $A = A^T$.

If A is real symmetric, then the eigenvectors corresponding to different eigenvalues are orthogonal.

Hence if A is an $n \times n$ real symmetric matrix with n distinct eigenvalues, then A is diagonalizable by an orthogonal matrix.

Theorem: For an $n \times n$ real symmetric matrix A:
(1) All the eigenvalues of A are real;
(2) A has n linearly independent eigenvectors.
(So such A is *always* diagonalizable by an orthogonal matrix.)

Let $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$ (see slides 5, 13).
Have eigenvalues $-3, -1, -4$, and corresponding eigenvectors $\boldsymbol{v}_1 = (1\ 0\ -1)^T$, $\boldsymbol{v}_2 = (1\ 2\ 1)^T$, $\boldsymbol{v}_3 = (1\ -1\ 1)^T$. Note A is real symmetric.
So (previous slide) $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ and $\boldsymbol{v}_3$ should be pairwise orthogonal:

10. Computation of eigenvalues

Some bounds on the eigenvalues of matrix A are given by:

1. If λ is an eigenvalue of A, then $|\lambda| \leq \|A\|$, where $\|\cdot\|$ is any of the three usual matrix norms.

2. (Gershgorin's Theorem)
Let $A = [a_{ij}]$ ne $n \times n$. Let D_i (for $i = 1, \ldots, n$) be the disc in the complex plane,

$$D_i = \left\{ z \in \mathbb{C} \,\middle|\, |z - a_{ii}| \leq \sum_{j=1, j\neq i}^{n} |a_{ij}| \right\}.$$

Every eigenvalue of A lies in at least one of these discs. And the union of any k of these discs which does not overlap the remaining $n-k$ discs contains exactly k (counting multiplicities) eigenvalues of A.

Examples

Take $A = \begin{pmatrix} -3 & 1 & 0 \\ 1 & -2 & 1 \\ 0 & 1 & -3 \end{pmatrix}$;
as before; eigenvalues are $-3, -1, -4$.
Discs are $D_1 = \{z \mid |z+3| \leq 1\}$,
$D_2 = \{z \mid |z+2| \leq 2\}$, and $D_3 = \{z \mid |z+3| \leq 1\}$
(so $D_1 = D_3$).

Let $B = \begin{pmatrix} 4 & -1 & 1 \\ 0 & 7 & -2 \\ 1 & 0 & -5 \end{pmatrix}$. Discs are
$D_1 = \{z \mid |z-4| \leq 2\}$, $D_2 = \{z \mid |z-7| \leq 2\}$, and $D_3 = \{z \mid |z+5| \leq 1\}$.

applying Gershgorin's theorem to the rows of A^T means using column sums from A. So there is a similar result for column discs from A.

See B above again.

3. The power method

Def'n: The eigenvalue λ of the matrix A is called the *dominant eigenvalue* of A if $|\lambda| > |\lambda'|$ for all eigenvalues $\lambda' \neq \lambda$ of A.
An eigenvalues corresponding to λ is a *dominant eigenvector*.

The Power Method gives a way of finding an approximate value of the dominant eigenvalue, and an approximation to a dominant eigenvector.

Form a sequence of vectors $\boldsymbol{u}_0, \boldsymbol{u}_1, \ldots, \boldsymbol{u}_k, \ldots$ where $\boldsymbol{u}_0$ is arbitrary, $\boldsymbol{u}_{k+1} = A\boldsymbol{u}_k$ (for $k \geq 0$). Then (usually) for k large,

(i) $\lambda \approx \dfrac{(\boldsymbol{u}_{k+1})_j}{(\boldsymbol{u}_k)_j}$, any $j \leq n$ with $(\boldsymbol{u}_k)_j \neq 0$

(ii) $\boldsymbol{u}_k \approx$ dominant eigenvector.

Example For $A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$, find exact value of dominant eigenvalue and eigenvector. Then try this approximation.

Usually get best estimate by taking the component j such that $|(\boldsymbol{u}_k)_j|$ is largest.

Anoth estimate for the dominant eigenvalue is

$$\lambda = \frac{\boldsymbol{u}_{k+1} \cdot \boldsymbol{u}_k}{\boldsymbol{u}_k \cdot \boldsymbol{u}_k} \quad \text{(dot products).}$$

This ratio is called a *Rayleigh quotient*.

Note: Power method gives only the dominant eigenvalue. For symmetric matrices, we can find the next most dominant one by *deflation*, based on:
Fact: If A is $n \times n$ with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ and $\boldsymbol{v}_1$ is an eigenvector corresponding to λ_1, then:

(a) $B = A - \left(\dfrac{\lambda_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right) \boldsymbol{v}_1 \boldsymbol{v}_1^T$ has eigenvalues $0, \lambda_2, \ldots, \lambda_n$.
(Note that $\boldsymbol{v}_1 \boldsymbol{v}_1^T$ is an $n \times n$ matrix.)

(b) If A is symmetric and $\boldsymbol{v}$ is an eigenvector of B corresponding to $\lambda_i \neq 0$, then $\boldsymbol{v}$ is also an eigenvector of A corresponding to λ_i.

So for symmetric A with eigenvalues $\lambda_1, \ldots, \lambda_n$ where $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n|$, having used the power method to find an approx'n to λ_1 and $\boldsymbol{v}_1$, form $B = A - \left(\dfrac{\lambda_1}{\boldsymbol{v}_1 \cdot \boldsymbol{v}_1}\right) \boldsymbol{v}_1 \boldsymbol{v}_1^T$ and repeat the power method on B to find an approx'n for λ_2 and $\boldsymbol{v}_2$.

In theory, could continue for λ_3 and $\boldsymbol{v}_3$, but in practise, round off error soon stops you.

Example Take $A = \begin{pmatrix} 3 & 2 \\ -1 & 0 \end{pmatrix}$ again.