Revision from week 17: a few statements about eigenvalues

Suppose $A\underline{v}_i = \lambda_i \underline{v}_i$, $\underline{v}_i \ne \underline{0}$, $i = 1, \ldots, n$.

Let $V = (\underline{v}_1, \ \ldots, \ \underline{v}_n)$ and let $D = \mathrm{diag}\{\lambda_1, \ldots, \lambda_n\}$.

$$AV = VD.$$

When the eigenvectors are linearly independent V has an inverse and A is **diagonalisable**. Otherwise A is said to be **non-diagonalisable** or **deficient**.

$\lambda_1, \ldots, \lambda_n$ is called the **spectrum** of A.

$$\rho(A) = \max\{|\lambda_1|, \ldots, |\lambda_n|\} = \textbf{spectral radius} \text{ of } A.$$

Vector norm axioms

$\|\underline{x}\| \ge 0 \ \forall \underline{x} \in \mathbb{R}^n$ with $\|\underline{x}\| = 0$ if and only if $\underline{x} = \underline{0}$.

$\|\alpha \underline{x}\| = |\alpha| \, \|\underline{x}\| \ \forall \alpha \in \mathbb{R}$ and $\forall \underline{x} \in \mathbb{R}^n$.

$\|\underline{x} + \underline{y}\| \le \|\underline{x}\| + \|\underline{y}\| \ \forall \underline{x}, \underline{y} \in \mathbb{R}^n$.

Matrix norm

The matrix norm induced by a vector norm is

$$\|A\| = \max\{\|A\underline{x}\| : \ \|\underline{x}\| = 1\}.$$

All of the following norm requirements are satisfied.

$\|A\| \ge 0 \ \forall A \in \mathbb{R}^{n,n}$ with $\|A\| = 0$ if and only if $A = 0$.

$\|\alpha A\| = |\alpha| \, \|A\| \ \forall \alpha \in \mathbb{R}$ and $\forall A \in \mathbb{R}^{n,n}$.

$\|A + B\| \le \|A\| + \|B\| \ \forall A, B \in \mathbb{R}^{n,n}$.

We also have $\|AB\| \le \|A\| \, \|B\|$.

The common vector norms

$$\begin{aligned}
\|\underline{x}\|_2 &= \left(x_1^2 + x_2^2 + \cdots + x_n^2\right)^{1/2} = \left(\underline{x}^T \underline{x}\right)^{1/2}, \\
\|\underline{x}\|_\infty &= \max\{|x_1|, \ldots, |x_n|\}, \\
\|\underline{x}\|_1 &= |x_1| + |x_2| + \cdots + |x_n|.
\end{aligned}$$

Expressions for the common matrix norms

$$\begin{aligned}
\|A\|_\infty &= \max_{1 \le i \le n} \sum_{j=1}^n |a_{ij}|, && \text{involves rows,} \\
\|A\|_1 &= \max_{1 \le j \le n} \sum_{i=1}^n |a_{ij}|, && \text{involves columns,} \\
\|A\|_2 &= \left(\rho(A^T A)\right)^{1/2} && \text{involves eigenvalues.}
\end{aligned}$$

For all these norms $\rho(A) \le \|A\|$. If $A^T = A$ then $\|A\|_2 = \rho(A)$.

Which $\underline{x}$ with $\|\underline{x}\| = 1$ gives the maximum?

∞-norm case:

$$(A\underline{x})_i = \sum_{j=1}^n a_{ij} x_j = \sum_{j=1}^n |a_{ij}|$$

If we choose $\underline{x} = (x_i)$ such that

$$a_{ij} x_j = |a_{ij}|, \quad j = 1, \ldots, n,$$

then $(A\underline{x})_i$ is a row sum of the absolute values. We do this for the row of A which gives the largest row sum.

1-norm case: It is one of the base vectors $\underline{e}_j$.

2-norm case: It is a normalised eigenvector of the symmetric matrix $A^T A$ corresponding to the largest eigenvalue of $A^T A$.

The matrix condition number

$$\kappa(A) = \|A\| \, \|A^{-1}\|, \quad 1 \le \kappa(A) \le \infty.$$

We say $\kappa(A) = \infty$ when A does not have an inverse.

$\kappa(A)$ is large when A is near to a matrix which has no inverse.

In the real symmetric case $\exists$ real eigenvalues λ_i and orthonormal eigenvectors $\underline{v}_i$. Let $V = (\underline{v}_1, \dots, \underline{v}_n)$, $D = \operatorname{diag}\{\lambda_i\}$ with

$$0 < |\lambda_n| \le \cdots \le |\lambda_1|.$$

In terms of V and D we have $A = VDV^T$ ($V^{-1} = V^T$ when V is orthogonal) which can be expressed in the form

$$A = \lambda_1 \underline{v}_1 \, \underline{v}_1^T + \cdots + \lambda_{n-1} \underline{v}_{n-1} \, \underline{v}_{n-1}^T + \lambda_n \underline{v}_n \, \underline{v}_n^T.$$

The nearest matrix to A in the 2-norm which is not invertible is

$$B = \lambda_1 \underline{v}_1 \, \underline{v}_1^T + \cdots + \lambda_{n-1} \underline{v}_{n-1} \, \underline{v}_{n-1}^T.$$

We can show that

$$\|A\|_2 = |\lambda_1|, \quad \|A - B\|_2 = |\lambda_n|, \quad \frac{\|A-B\|_2}{\|A\|_2} = \left|\frac{\lambda_n}{\lambda_1}\right| = \frac{1}{\kappa(A)}.$$

The matrix condition number continued

$$\kappa(A) = \|A\| \, \|A^{-1}\|, \quad 1 \le \kappa(A) \le \infty.$$

- Do we compute it?

  Generally no but we might estimate it.

- What is it used for in this module?

  It quantifies when a matrix is nearly singular and for the problem

  $$A\underline{x} = \underline{b}$$

  it is such that if we change the entries of A or $\underline{b}$ by terms of size ϵ then the solution may change by magnitude of about $\kappa(A)\epsilon$. It helps quantify the sensitivity of the system to changes to A and/or $\underline{b}$.
- Note that the ratio of the extreme eigenvalues only describes the condition number when A is real and symmetric.

Chap 2: Direct methods for solving $A\underline{x} = \underline{b}$

1. How would you solve the following 6×6 linear system?

$$\begin{pmatrix} 3 & 0 & 1 & 2 & 4 & 6 \\ 1 & 7 & 9 & 2 & 2 & 0 \\ 4 & 5 & 9 & 8 & 2 & 1 \\ 3 & 3 & 3 & 1 & 1 & 1 \\ 8 & 4 & 0 & 3 & 5 & 2 \\ 4 & 7 & 9 & 8 & 6 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \\ x_6 \end{pmatrix} = \begin{pmatrix} -15 \\ 30 \\ 46 \\ 6 \\ -3 \\ 38 \end{pmatrix}$$

   A very small problem for a computer but a bit tiring to attempt to do by hand calculations.

2. If A is $n \times n$ with $n = 8000$ then how long does it take to solve $A\underline{x} = \underline{b}$ on a computer? Are the methods reliable and accurate?

Using Matlab

1. We get the answer in the 6×6 case by putting the following.

```
A=[3 0 1 2 4 6
   1 7 9 2 2 0
   4 5 9 8 2 1
   3 3 3 1 1 1
   8 4 0 3 5 2
   4 7 9 8 6 3];
b=[-15 30 46 6 -3 38]';

x=A\b
```

2. In this part of this module we describe the Gauss elimination/ LU factorization method that is usually used. The number of operations grows with n like n^3 for a full matrix. With $n = 8000$ I have a timing of about 6 seconds on a laptop new in 2015. With $n = 16000$ it took about 47 seconds. In practice the method is reliable but it is not guaranteed to work in every case.

Upper triangular systems

$$\begin{array}{cccccccc}
u_{11}x_1 & + & u_{12}x_2 & + & \cdots & + & u_{1n}x_n & = & b_1 \\
 & + & u_{22}x_2 & + & \cdots & + & u_{2n}x_n & = & b_2 \\
 & & \ddots & & & & & & \vdots \\
 & & & & u_{n-1,n-1}x_{n-1} & + & u_{n-1,n}x_n & = & b_{n-1} \\
 & & & & & & u_{nn}x_n & = & b_n
\end{array}$$

Backward substitution:

$$\begin{aligned}
x_n &= b_n/u_{nn} \\
x_i &= \left(b_i - \sum_{k=i+1}^{n} u_{ik}x_k \right) /u_{ii}, \quad i = n-1, \ldots, 1 \ .
\end{aligned}$$

$\mathcal{O}(n^2/2)$ entries in U and about n^2 operations to get $\underline{x}$.

With lower triangular systems we use forward substitution which has the same number of operations.

Solving $A\underline{x} = \underline{b}$ when $A = LU$

Here L is lower triangular and U is upper triangular.

$$A\underline{x} = LU\underline{x} = \underline{b}.$$

Algorithm:

Solve $L\underline{y} = \underline{b}$ by forward substitution.

Solve $U\underline{x} = \underline{y}$ by backward substitution.

The number of operations is about the same as computing

$$A^{-1}\underline{b}$$

if the inverse matrix A^{-1} is available.

It is rare to need to have A^{-1}.

Some of the material on the remaining slides will probably be covered next week.

Reduction to triangular form

Basic reduction is a specific order of the operations. At each stage in the basic reduction process we create zeros below the diagonal in a column and we have a vector of the **multipliers**.

$$A = \begin{pmatrix} \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \\ \times & \times & \times & \times \end{pmatrix} \quad \to \quad \begin{pmatrix} \times & \times & \times & \times \\ 0 & \mathsf{X} & \mathsf{X} & \mathsf{X} \\ 0 & \mathsf{X} & \mathsf{X} & \mathsf{X} \\ 0 & \mathsf{X} & \mathsf{X} & \mathsf{X} \end{pmatrix} = A^{(1)}, \quad \underline{m}_1 = \begin{pmatrix} 0 \\ m_{21} \\ m_{31} \\ m_{41} \end{pmatrix},$$

$$\to \quad \begin{pmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \mathsf{X} & \mathsf{X} \\ 0 & 0 & \mathsf{X} & \mathsf{X} \end{pmatrix} = A^{(2)}, \quad \underline{m}_2 = \begin{pmatrix} 0 \\ 0 \\ m_{32} \\ m_{42} \end{pmatrix},$$

$$\to \quad \begin{pmatrix} \times & \times & \times & \times \\ 0 & \times & \times & \times \\ 0 & 0 & \times & \times \\ 0 & 0 & 0 & \mathsf{X} \end{pmatrix} = U, \quad \underline{m}_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \\ m_{43} \end{pmatrix}.$$

x and X are potentially non-zero.

The $A = LU$ factorization

When the basic reduction is possible we have the factorization

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ m_{21} & 1 & 0 & 0 \\ m_{31} & m_{32} & 1 & 0 \\ m_{41} & m_{42} & m_{43} & 1 \end{pmatrix} \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix}.$$

$$L = M_1^{-1}M_2^{-1}M_3^{-1} = I + \underline{m}_1\underline{e}_1^T + \underline{m}_2\underline{e}_2^T + \underline{m}_3\underline{e}_3^T,$$

where each $M_k = I - \underline{m}_k\underline{e}_k^T$ is a **Gauss transformation matrix**.

Later we write $l_{ij} = m_{ij}$ for the entries of the lower triangular matrix. As the diagonal entries of L are all equal to 1 the matrix is said to be **unit lower triangular**.