Question: are the eigenvalues of A and A^T the same? yes

are the eigenvectors of A and A^T the same? maybe not

Let x be an eigenvector of $A \in \mathbb{R}^{n\times n}$ with eigenvalue λ.

Claim that λ is also an eigenvalue of A^T.

$$Ax = \lambda x$$

$$Ax \cdot y = \lambda x \cdot y \qquad \text{for any } y \in \mathbb{R}^n$$

$$x \cdot A^T y = x \cdot \lambda y$$

$$x \cdot A^T y - x \cdot \lambda y = 0$$

$$x \cdot (A^T - \lambda I) y = 0 \qquad \text{for any } y \in \mathbb{R}^n$$

If λ is an eigenvalue of A^T there is nothing to do because that is what we're trying to show.

Suppose, for contradiction, that λ is not an eigenvalue of A^T, $\det(A^T - \lambda I) \neq 0$ or in other words that $A^T - \lambda I$ is invertible.

Thus $(A^T - \lambda I)^{-1}$ exists.

Define $y = (A^T - \lambda I)^{-1} x$ and plug this value for y in

$$x \cdot \cancel{(A^T - \lambda I)}\cancel{(A^T - \lambda I)^{-1}} x = 0$$

Therefore $x \cdot x = 0$ or $\|x\|_2 = 0$ or $x = 0$.

Since x is an eigenvector it can't be zero. Therefore λ must be an eigenvalue of A^T

We'll come back to eigenvalues and eigenvectors

5 Eigenvalues and eigenvectors of a symmetric matrix

5.1 Introduction

But first:

4 Simultaneous nonlinear equations

Note: we are skipping Chapter 3.

First day of class idea to solve $f(x)=0$ for x...

1-step iteration → Newton's method.

$f: \mathbb{R} \to \mathbb{R}$

also secant method ← 2-step iteration

focus on these two

→ relaxation method

and bisection method...

too big too small doesn't make sense for vectors

Now suppose $f: \mathbb{R}^n \to \mathbb{R}^n$ so $f(x)=0$ represents solving a system of n equations in n unknowns...

Any 1-step method is of the form

$$x^{(k+1)} = g\left(x^{(k)}\right)$$ where $x^{(0)}$ is an initial approx.

Relaxation method

easy to write, but does it work? for what choice of λ?

$$g(x) = x - \lambda f(x)$$

scalar multiplier... (relaxation coefficient)

Newton's method

Idea to make a linear approximation and solve the linear approximation equal to 0.

$$f(x) \approx \underbrace{f(x_0) + Df(x_0)(x - x_0)}$$

linear approximation of f at x_0 is actually the definition of the derivative in higher dimensions.

Math 283 ...

$$J_f(x_0) = Df(x_0) = \left.\begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}\right|_{x=x_0}$$

notation in book.

Solve for x in $f(x_0) + Df(x_0)(x - x_0) = 0$

$$Df(x_0)(x - x_0) = -f(x_0)$$

$$\cancel{Df(x_0)^{-1}}\, \cancel{Df(x_0)}(x - x_0) = Df(x_0)^{-1}\left(-f(x_0)\right)$$

$$x - x_0 = -Df(x_0)^{-1} f(x_0)$$

Therefore $\quad x = x_0 - Df(x_0)^{-1} f(x_0)$

Therefore the iteration is

$$x^{(k+1)} = x^{(k)} - \left(Df(x^{(k)})\right)^{-1} f(x^{(k)})$$

or $\quad g(x) = x - \left(Df(x)\right)^{-1} f(x)$

$(J_f(x))^{-1}$ in the book

Question if $Df(\xi)$ is invertible for the root $f(\xi) = 0$ and the second derivative of f exists. Then does our theory about quadratic convergence still work?

Idea: use multi-variable form of Taylor's theorem.

Compare to relaxation:

$$g(x) = x - \lambda f(x)$$

A natural compromise between relaxation an Newton's method with n equations/unknowns is

$$g(x) = x - Mf(x)$$

$M \in \mathbb{R}^{n\times n}$

and likely $M \approx (Df(\xi))^{-1}$

In order for $x^{(k+1)} = g(x^{(k)})$ to converge, the approximations need to get closer together with each iteration. That because if $x^{(k)}$'s are getting closer to the solution ξ then they must be getting closer to each other...

That means a contraction

$$\|x^{(k+2)} - x^{(k+1)}\| \leq L\|x^{(k+1)} - x^{(k)}\| \quad \text{for some } L \in (0,1)$$

$$\|g(x^{(k+1)}) - g(x^{(k)})\| \leq L\|x^{(k+1)} - x^{(k)}\| \quad \text{for some } L \in (0,1)$$

In general we want

$$\|g(x) - g(y)\| \leq L\|x - y\| \quad \text{for some } L \in (0,1).$$

(this is a type of Lipschitz function)

Note that g must be continuous.

$$\underbrace{\|g(x) - g(y)\|}_{\varepsilon} \leq L\underbrace{\|x - y\|}_{\delta}$$

... can ensure $\|g(x) - g(y)\| < \varepsilon$ using $\|x - y\| < \delta = \varepsilon/L$

Consecutive iteration being close is good but what about

$\| x^{(m)} - x^{(k)} \|$ How close is $x^{(m)}$ from $x^{(k)}$?

Assume $m > k$.

$$x^{(m)} - x^{(k)} = x^{(m)} - x^{(m-1)} + x^{(m-1)} - x^{(m-2)} + \cdots + x^{(k-1)} - x^{(k)}$$

If any two iterates get closer together as $m \to \infty$ and $k \to \infty$ we can infer that the limit

$x^{(k)} \to \xi$ exists

and continuity then implies $g(\xi) = \xi$.