Lecture 2: Linear Algebra: Eigenvalues, systems of differential equations, etc., S. Ch 1

Main problems of Numerical Linear Algebra

Find eigenvector(s) **x** and complex eigenvalue(s) λ
"Standard" eigenvalue problem:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

Generalized eigenvalue problem:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{M}\mathbf{x}$$

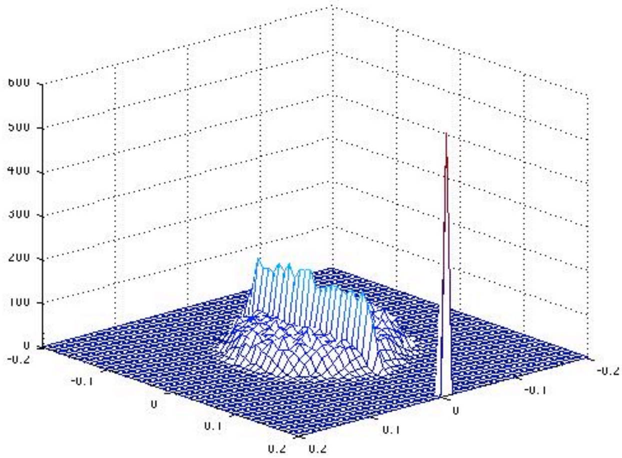
```
Matlab: lam = eig(A);
```

Examples
1. The Markov chain equilibrium distribution - Lect 1.
Uncommon that

- the eigenvalue is known
- the eigenvalue with largest absolute value is wanted

Here is a plot of a computational experiment on the distribution of complex eigenvalues to 600 40x40 Markov random matrices, i.e., a 2D histogram of 600 x 40 = 2400 points.

1. the 600 peak at the extreme right (not at the right spot ...)
2. flat distribution around 0 with a radius of ≈ 0.1
3. a ridge of real eigenvalues



I think 1. is easy and can think of a reason for 3., but 2. ?? One would think that since **A** has only positive entries the eigenvalues would flock to the right hand plane. *Not so.* You may want to check on Alan Edelman's lectures on random matrices.
While on the subject of guessing eigenvalues, look at the Gershgorin circle theorem (p. 570):
Every eigenvalue is in the union of circles C_i, i = 1,2,...,n

$$C_i : |\lambda - a_{ii}| \le \sum_{j \ne i} |a_{ij}|$$

The G.-circles for a Markov matrix are all centered on the interval [0,1] and pass through 1, so all contained in the unit circle. But the probable eigenvalues occupy but a minuscule portion of it.

2. Quadratic forms $q(\mathbf{x}) = \mathbf{x}^T\mathbf{K}\mathbf{x}$

Note: Uses only the symmetric part $\frac{1}{2}(\mathbf{K}+\mathbf{K}^T)$ so consider $\mathbf{K}$ symmetric.
We will measure the size of $\mathbf{x}$ by its Euclidean norm,

$$\|\mathbf{x}\|_2 = \left(\sum |x_k|^2\right)^{1/2} = \sqrt{\mathbf{x}^T\mathbf{x}}$$

and remind you of the triangle inequality, the multiplication by scalar, the Cauchy-Schwarz inequality

$$\mathbf{x}^T\mathbf{y} \le \|\mathbf{x}\|_2 \cdot \|\mathbf{y}\|_2$$

and the definition of the operator norm of a linear operator (matrix!) $\|\mathbf{A}\|_2$, induced by the vector norm

$$\|\mathbf{A}\|_2 = \max \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2} = \max \sqrt{\frac{\mathbf{x}^T\mathbf{A}^T\mathbf{A}\mathbf{x}}{\mathbf{x}^T\mathbf{x}}},$$

and the Rayleigh quotient (p 219)

$$R_{\mathbf{K}}(x) = \mathbf{x}^T\mathbf{K}\mathbf{x}/\mathbf{x}^T\mathbf{x}$$

The norm shows the maximal magnification possible in the mapping. Let us compute it by finding the max. of the Rayleigh quotient by differentiation:

$$\frac{\partial}{\partial x_k}\left(\mathbf{x}^T\mathbf{K}\mathbf{x}\right) = \frac{\partial}{\partial x_k}\left(\sum_{i,j} k_{ij}x_i x_j\right) = \sum_{i,j} k_{ij}(\underbrace{\frac{\partial x_i}{\partial x_k}}_{=\delta_{ik}} x_j + x_i \underbrace{\frac{\partial x_j}{\partial x_k}}_{=\delta_{jk}}) =$$

$$= \sum_j k_{kj}x_j + \sum_i k_{ik}x_i = 2\sum_j k_{kj}x_j = 2(\mathbf{K}\mathbf{x})_k$$

so $\partial R/\partial x_k = 0$, $k = 1,...,n$, if (and only if)

$$2\mathbf{K}\mathbf{x}\cdot(\mathbf{x}^T\mathbf{x}) - (\mathbf{x}^T\mathbf{A}\mathbf{x})\cdot 2\mathbf{x} = 0$$

or

$$\mathbf{K}\mathbf{x} = \underbrace{R_{\mathbf{K}}(\mathbf{x})}_{\lambda}\mathbf{x}$$

Theorem:

If $\mathbf{K}\mathbf{x}^* = \lambda\mathbf{x}^*$, then $\mathbf{x}^*$ is a stationary point of $R_{\mathbf{K}}(\mathbf{x})$ and $\lambda = R_{\mathbf{K}}(\mathbf{x}^*)$, and conversely.

It follows that

$$\|\mathbf{A}\|_2 = \sqrt{\text{largest eigenvalue of } \mathbf{A}^T\mathbf{A}}$$

Note: minimization of quadratic forms with a single quadratic constraint also leads to eigenvalue problems.

3: Linear differential equations with constant coefficients
(S p 53, Ch 2.1, 2.2)
Basics:
Exponentials are eigenfunctions of differential and difference operators with constant coefficients. Usually the independent variable is now time t.
Let $D = d/dt$. Then $D\exp(\lambda t) = \lambda \exp(\lambda t)$, and

$$p(D)e^{\lambda t} = \sum_{k=1}^{n} a_k D^k (e^{\lambda t}) = \sum_{k=1}^{n} a_k \lambda^k (e^{\lambda t}) = p(\lambda)e^{\lambda t}$$

The analogue holds for difference operators (p 54 ff), e.g.

$$\Delta u(t_k) = u(t_{k+1}) - u(t_k),\ t_{k+1} - t_k = h,\ k = ...,-1,0,1,...$$

Then

$$\Delta e^{\lambda t} = e^{\lambda(t+h)} - e^{\lambda t} = \mu e^{\lambda t}, \mu = e^{\lambda h} - 1$$

Note that $\lim_{h \to 0} \mu / h = \lambda$

$$p(\Delta)e^{\lambda t} = \sum_{k=1}^{n} a_k \Delta^k (e^{\lambda t}) = \sum_{k=1}^{n} a_k \mu^k (e^{\lambda t}) = p(\mu)e^{\lambda t}$$

An example with complex eigenvalues, rigid body rotation in a plane:
The velocity at (x,y) of rotation with angular velocity ω around the origin is

$$(dx/dt, dy/dt) = \omega(-y,x)$$

or

$$\frac{d}{dt}\mathbf{u} = \underbrace{\omega\begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}}_{\mathbf{A}}\mathbf{u}, \mathbf{u}(t) = \begin{pmatrix} x(t) \\ y(t) \end{pmatrix}$$

Look for special solution vectors $\mathbf{u} = \mathbf{v}\exp(\lambda t)$, $\mathbf{v}$ some constant vector:

$$\frac{d}{dt}\mathbf{u}(t) = \lambda e^{\lambda t}\mathbf{v} = \mathbf{Au} = e^{\lambda t}\mathbf{v} \Leftrightarrow \lambda\mathbf{v} = \mathbf{Av}$$

The eigenvalues of $\mathbf{A}$ are imaginary, $+/-\omega i$ and the eigenvetors are $(-1,i)^T$ and $(1,i)^T$ so any linear combination

$$\mathbf{u}(t) = a\begin{pmatrix} -1 \\ i \end{pmatrix}e^{+i\omega t} + b\begin{pmatrix} 1 \\ i \end{pmatrix}e^{-i\omega t}$$

satisfies the equation. There are enough (2) integration constants to satisfy initial conditions like $\mathbf{u}(0) = (x0,y0)$, so this is the general solution. It looks complex, but real initial conditions fix that:

$$\mathbf{u}(t) = \frac{-x0 - iy0}{2}\begin{pmatrix} -1 \\ i \end{pmatrix}e^{+i\omega t} + \frac{x0 - iy0}{2}\begin{pmatrix} 1 \\ i \end{pmatrix}e^{-i\omega t} = \begin{pmatrix} x0 \cdot \cos\omega t - y0 \cdot \sin\omega t \\ x0 \cdot \sin\omega t + y0 \cdot \cos\omega t \end{pmatrix}$$

$$= \underbrace{\begin{pmatrix} \cos\omega t & -\sin\omega t \\ \sin\omega t & \cos\omega t \end{pmatrix}}_{\text{Rotates angle } \omega t}\begin{pmatrix} x0 \\ y0 \end{pmatrix}$$

The manipulations become trivial if we use complex variables: $z(t) = x(t) + iy(t)$.
Then

$dz/dt = i\omega z$ and $z(t) = \exp(i\omega t)z(0)$

so

$\text{angle}(z(t)) = \text{angle}(z(0)) + \omega t$, $|z(t)|=|z(0)|$

You may want to think about the 3D counterpart, rotation with angular velocity ω around a unit length vector (w1,w2,w3)

$$\frac{d}{dt}\mathbf{u}(t) = \omega\mathbf{w}\times\mathbf{u} = \mathbf{A}\mathbf{u}, \mathbf{A} = \omega\begin{pmatrix} 0 & -w3 & w2 \\ w3 & 0 & -w1 \\ -w2 & w1 & 0 \end{pmatrix}$$

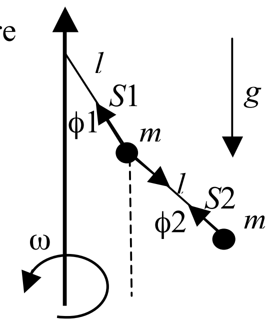
Compute the eigenvalues of **A**! Hint: One is zero ... compute det**A** to see this.

4: Bifurcation ... linearized

(see also S. p 108-109)

Consider a double pendulum. Its motion is constrained by the joints to a plane, rotating around a vertical axis with angular velocity ω. Compute its equilibrium position, assuming it has small angles ϕ_1, ϕ_2!

The equilibrium equations for the mass points 1 and 2 are



$$\begin{aligned} ml\omega^2 \sin\phi_1 + S_2 \sin\phi_2 &= S_1 \sin\phi_1 \\ S_1 \cos\phi_1 &= S_2 \cos\phi_2 + mg \\ ml\omega^2 (\sin\phi_1 + \sin\phi_2) &= S_2 \sin\phi_2 \\ S_2 \cos\phi_2 &= mg \end{aligned}$$

Eliminate the forces Si:

$$S_1 = 2mg/\cos\phi_1$$
$$S_2 = mg/\cos\phi_2 :\Rightarrow$$
$$\begin{cases} (\sin\phi_1 + \sin\phi_2)\lambda = \tan\phi_2 \\ \lambda\sin\phi_1 = 2\tan\phi_1 - \tan\phi_2 \end{cases}$$
$$\lambda = \frac{\omega^2 l}{g}$$

Approximate the trig-functions to produce the final linear system:

$$\begin{cases} (\phi_1 + \phi_2)\lambda = \phi_2 \\ \lambda\phi_1 = 2\phi_1 - \phi_2 \end{cases} \Rightarrow \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix} = \lambda\begin{pmatrix} \phi_1 \\ \phi_2 \end{pmatrix}$$

which shows that: unless λ is an eigenvalue, there is only the trivial solution: the pendulum hangs straight down. For sufficiently small ω, λ < 2- sqrt(2), this is the case. The two eigenvectors are (1,sqrt(2)) and (1,-sqrt(2)). Both shapes can be provoked when you twirl a hanging rope. But only the non-linear model tells what happens after the first bifurcation when ω is increased. The story requires that we consider the time-dependent problem, later.