Lecture 23: Eigenstuff

An Example (population dynamics)

We begin with a linear system of difference equations (or "discrete dynamical system")

$$W(\tau+1) = 0.86 \cdot W(\tau) + 0.08 \cdot S(\tau)$$

$$S(\tau+1) = -0.12 \cdot W(\tau) + 1.14 \cdot S(\tau)$$

describing the wolf & sheep populations over time (measured in months). Set

$$\vec{x}(\tau) = \begin{pmatrix} W(\tau) \\ S(\tau) \end{pmatrix}, \quad A = \begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix},$$

so that

$$\vec{x}(\tau+1) = A \cdot \vec{x}(\tau).$$

To understand the system's long-term behavior, we need to compute

$$\vec{x}(\tau) = A^\tau \vec{x}_0 = \underbrace{A \cdot \cdots \cdot A}_{\tau \text{ times}} \begin{pmatrix} W_0 \\ S_0 \end{pmatrix}$$

← starting values at time $\tau = 0$

for very large τ.

- Consider the initial state vector $W_0 = 100$, $S_0 = 300$

  $$\vec{x}(1) = A\vec{x}_0 = \begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix}\begin{pmatrix} 100 \\ 300 \end{pmatrix} = \begin{pmatrix} 110 \\ 330 \end{pmatrix} = 1.1 \cdot \vec{x}_0$$

  Then $\vec{x}(\tau) = A^\tau \vec{x}_0 = (1.1)^\tau \vec{x}_0 \implies$ populations grow together.

- Consider the initial state vector $W_0 = 200$, $S_0 = 100$.

$$\vec{x}(1) = A\vec{x}_0 = \begin{pmatrix} 0.86 & 0.08 \\ -0.12 & 1.14 \end{pmatrix}\begin{pmatrix} 200 \\ 100 \end{pmatrix} = \begin{pmatrix} 180 \\ 90 \end{pmatrix} = 0.9\,\vec{x}_0$$

and $\vec{x}(\tau) = A^\tau \vec{x}_0 = (0.9)^\tau \vec{x}_0 \implies$ populations shrink together: too many wolves eating too few sheep.

Put $\vec{v}_1 = \begin{pmatrix} 100 \\ 300 \end{pmatrix}$, $\vec{v}_2 = \begin{pmatrix} 200 \\ 100 \end{pmatrix}$ → "eigenvectors"

$\lambda_1 = 1.1$, $\lambda_2 = 0.9$. → "eigenvalues"

Now, in your backyard, $\vec{x}_0 = \begin{pmatrix} 1000 \\ 1000 \end{pmatrix}$. So

$$\vec{x}(1) = A\vec{x}_0 = \begin{pmatrix} 940 \\ 1020 \end{pmatrix},$$

and we aren't in one of the neat cases above. But we can still use those cases to understand what is going on in the long term: break $\vec{x}_0 = c_1\vec{v}_1 + c_2\vec{v}_2 \implies$

$$\vec{x}(1) = A\vec{x}_0 = c_1 A\vec{v}_1 + c_2 A\vec{v}_2 = c_1(1.1)\vec{v}_1 + c_2(0.9)\vec{v}_2$$

and $$\vec{x}(\tau) = \underbrace{c_1(1.1)^\tau \vec{v}_1}_{\text{grows by 10\% every month}} + \underbrace{c_2(0.9)^\tau \vec{v}_2}_{\text{shrinks by 10\% every month}}.$$

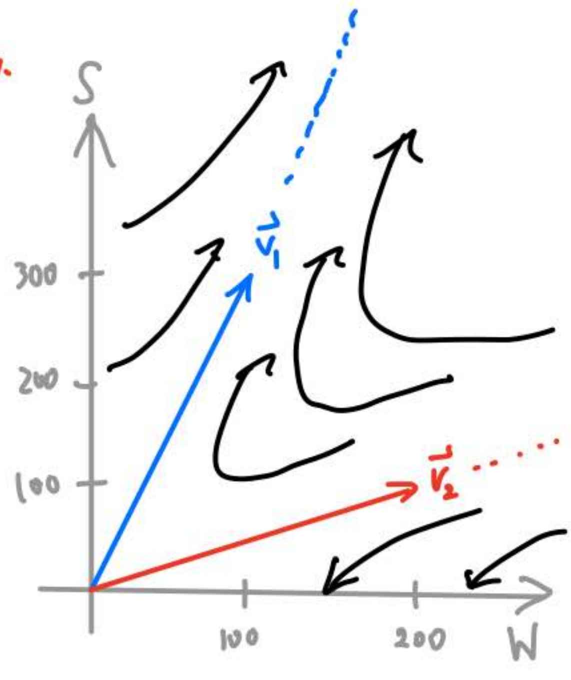
So in case 4

$$\vec{x}_0 = \begin{pmatrix} 1000 \\ 1000 \end{pmatrix} = 2\begin{pmatrix} 100 \\ 300 \end{pmatrix} + 4\begin{pmatrix} 200 \\ 100 \end{pmatrix} = 2\vec{v}_1 + 4\vec{v}_2,$$

we have

$$\vec{x}(\tau) = 2(1.1)^\tau \begin{pmatrix} 100 \\ 300 \end{pmatrix} + 4(0.9)^\tau \begin{pmatrix} 200 \\ 100 \end{pmatrix}.$$

The flow lines at right depict the path taken by $\vec{x}(\tau)$. Clearly we need to get the sheep/wolf population ratio below 1:2 . . .



$\boxed{\text{A little theory}}$

Let A be an $n \times n$ matrix.

Definition: A nonzero vector $\vec{v} \in \mathbb{R}^n$ is called an eigenvector of A if

$$A\vec{v} = \lambda \vec{v}$$

for some scalar λ. In this case, λ is called an eigenvalue of A, and $\vec{v}$ is called an eigenvector (of A) with eigenvalue λ.

Let

$$E_\lambda := \{\vec{v} \in \mathbb{R}^n \mid A\vec{v} = \lambda \vec{v}\}$$

be the set of all eigenvectors with eigenvalue λ (together with the zero vector). This is the eigenspace associated to λ. Now we make the fundamental

Observation: $A\vec{v} = \lambda\vec{v} \iff (A - \lambda I_n)\vec{v} = \vec{0}$,

i.e. $$E_\lambda = \mathrm{Nul}(A - \lambda I_n),$$ and is therefore a subspace of $\mathbb{R}^n$! In addition,

$$\begin{aligned} \lambda \text{ is an eigenvalue of } A &\iff \mathrm{Nul}(A - \lambda I_n) \neq \{\vec{0}\} \\ &\iff A - \lambda I_n \text{ is \underline{not} invertible} \\ &\iff \det(A - \lambda I_n) = 0, \end{aligned}$$

proving

Theorem 1: The eigenvalues of A are the solutions to the characteristic equation

$$\det(A - \lambda I_n) = 0.$$

Suppose A is upper triangular, say

$$A = \begin{pmatrix} \alpha_1 & * & * \\ 0 & \alpha_2 & * \\ 0 & 0 & \alpha_3 \end{pmatrix}.$$ [$*$ = anything]

Then $A - \lambda I_3 = \begin{pmatrix} \alpha_1 - \lambda & * & * \\ 0 & \alpha_2 - \lambda & * \\ 0 & 0 & \alpha_3 - \lambda \end{pmatrix}$

$$\Rightarrow \det(A - \lambda I_3) = (\alpha_1 - \lambda)(\alpha_2 - \lambda)(\alpha_3 - \lambda).$$

Corollary: The eigenvalues of an upper triangular matrix are the diagonal entries.

Another Example

Let $A = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$. Find all the eigenvalues and bases/dimension for each eigenspace.

Ⓐ First solve the characteristic equation

$$\det(A-\lambda I_3) = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 1 & 1 & 1-\lambda \end{vmatrix} \quad \text{subtract} \quad = \begin{vmatrix} 1-\lambda & 1 & 1 \\ 1 & 1-\lambda & 1 \\ 0 & \lambda & -\lambda \end{vmatrix}$$

add

$$= \begin{vmatrix} 1-\lambda & 1 & 2 \\ 1 & 1-\lambda & 2-\lambda \\ 0 & \lambda & 0 \end{vmatrix} = -\lambda \begin{vmatrix} 1-\lambda & 2 \\ 1 & 2-\lambda \end{vmatrix}$$

now do Laplace

$$= -\lambda\{(\lambda-1)(\lambda-2) - 2\cdot 1\} = -\lambda\{\lambda^2 - 3\lambda \cancel{+2-2}\}$$

$$= -\lambda^2(\lambda-3)$$

$\Longrightarrow$ eigenvalues are 0 & 3.

Ⓑ Then compute bases for E_0 & E_3, viewed as null spaces

- For $E_3 = \mathrm{Nul}(A - 3I_3) = \mathrm{Nul}\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$

row-reduce: $\begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{pmatrix} \to \begin{pmatrix} 1 & -2 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{pmatrix}$

x_3 free

$\Longrightarrow$ basis for E_3 is $\left\{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}\right\} \Longrightarrow \dim E_3 = 1$.

- For $E_0 = \mathrm{Nul}(A - 0I_3) = \mathrm{Nul}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix}$

row-reduce: $\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$ $\Longrightarrow$ basis for E_0 is $\left\{\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}\right\}$

x_2, x_3 free

$\Longrightarrow \dim E_0 = 2$.

And in fact, putting the two bases together gives a basis for $\mathbb{R}^3$. This doesn't always happen. In the next lecture we'll see why it does happen in some cases (including here).

$\boxed{\text{Some problems}}$

(1) Given 2 eigenvectors of A, is their sum an eigenvector?

(2) In each case, is $\vec{v}$ an eigenvector of A? If so, find the eigenvalue:

(a) $\vec{v} = \begin{pmatrix} 1 \\ 4 \end{pmatrix}$, $A = \begin{pmatrix} -3 & 1 \\ -3 & 8 \end{pmatrix}$; (b) $\vec{v} = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, $A = \begin{pmatrix} 4 & -1 \\ -1 & 4 \end{pmatrix}$

(3) Find the eigenvalues and eigenvectors of $A = \begin{pmatrix} 7 & 3 \\ 3 & -1 \end{pmatrix}$.

$\boxed{\text{Application : Google PageRank}}$

Number all of the webpages on the internet $1, \dots, N$.

Denote by ℓ_j the number of pages that the j^{th} page links to (we'll assume this isn't 0). Define

$$L_{ij} := \begin{cases} 1 & \text{if page } j \text{ links to page } i \\ 0 & \text{otherwise}. \end{cases}$$

If $p_i(m)$ is the probability that a random web surfer is on page i at time m, then

$$(*) \quad p_i(m+1) = \sum_{j=1}^{N} \frac{L_{ij}}{l_j} p_j(m).$$

Writing S for the matrix with $(i,j)^{\text{th}}$ entry L_{ij}/l_j, we can rewrite $(*)$ as

$$(**) \qquad \vec{p}(m+1) = S\,\vec{p}(m).$$

S is a stochastic matrix and $(**)$ is a Markov chain. It has an eigenvector $\vec{r}$ with eigenvalue 1, i.e. $S\vec{r} = \vec{r}$, and if we normalize $\vec{r}$ to have its entries sum to 1, then it is unique.

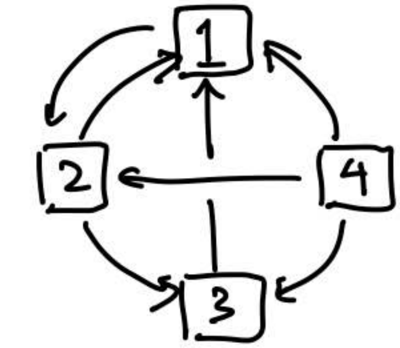
This steady-state vector is a simplification of the <u>PageRank vector</u> introduced by Sergey Brin and Larry Page. Think of it as $\lim_{m \to \infty} \vec{p}(m)$ (to be explained more in Lect. 24).

<u>Aside</u> The non-simplified version takes

$$\vec{p}(m+1) = \frac{1-d}{N} \cdot \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} + d \cdot S\,\vec{p}(m)$$

for some d (such as 0.85), which is the probability that the web surfer follows a link — as opposed to hopping to a completely random webpage (prob. $1-d$).

$\boxed{\text{Example}}$ The internet has 4 webpages, with links



$$\rightsquigarrow L = \begin{array}{cccc} \text{1 links to} & \text{2 links to} & \text{3 links to} & \text{4 links to} \\ \end{array}$$

$$L = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow S = \begin{pmatrix} 0 & 1/2 & 1 & 1/3 \\ 1 & 0 & 0 & 1/3 \\ 0 & 1/2 & 0 & 1/3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

We want to solve $S\vec{r} = \vec{r}$, or $(S - I)\vec{r} = \vec{0}$.

Row-reducing $S - I = \begin{pmatrix} -1 & 1/2 & 1 & 1/3 \\ 1 & -1 & 0 & 1/3 \\ 0 & 1/2 & -1 & 1/3 \\ 0 & 0 & 0 & -1 \end{pmatrix}$ gives $\begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$

whose null space is spanned by $\begin{pmatrix} 2 \\ 2 \\ 1 \\ 0 \end{pmatrix}$. So $\vec{r} = \begin{pmatrix} 2/5 \\ 2/5 \\ 1/5 \\ 0 \end{pmatrix}$.

If we use the modified version with $d = 0.9$,

we need to solve $\vec{r} = \frac{0.1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} + 0.9\, S\vec{r}$, which gives

$(I - 0.85\, S)\vec{r} = \frac{0.1}{4}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$ or

$$\vec{r} = \frac{0.1}{4}(I - 0.9 \cdot S)^{-1}\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \approx \begin{pmatrix} 0.389 \\ 0.382 \\ 0.204 \\ 0.025 \end{pmatrix}$$

$\uparrow$ entries add to 1

So the rank of the pages is indeed 1, 2, 3, then 4.