Mathematics for Economists: Lecture 12

Juuso Välimäki

Aalto University School of Business

Spring 2020

This lecture covers

1. Eigenvalues and eigenvectors of matrices
2. Diagonalizing a matrix
3. Non-diagonalizable matrices
4. Examples of systems of difference equations

Eigenvalues and eigenvectors of a matrix

- Consider next linear systems with constant coefficients.
- Let $x_t \in \mathbb{R}^n$ for all t and let A be an $n \times n$ matrix of real numbers.
- A linear homogenous system is given by:

$$x_{t+1} = Ax_t.$$

- As before, we can 'solve' this by repeated substitution to get

$$x_{t+k} = A^k x_t.$$

- Hence I could write the general solution as $x_t = A^t c$ for some vector $c = (c_1, ..., c_k)$.

Eigenvalues and eigenvectors of a matrix

- I do not consider this a real solution since it is almost impossible to see what A^t is except in some very special cases.
- If A is a diagonal matrix with diagonal elements $a_1, ..., a_n$, then the solution becomes

$$x_{i,t} = c_i a_i^t \text{ for } i \in \{1, ..., n\}.$$

- For example, if

$$\begin{pmatrix} x_{1,t+1} \\ x_{2,t+1} \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix},$$

  then we have $x_{1,t} = c_1 2^t, x_{2,t} = c_2 3^t$.
- In this example, x_i does not depend at all on x_j and therefore the two equations can be solved separately.

Eigenvalues and eigenvectors of a matrix

- In general, this is not so easy.
- We want to change the basis in $\mathbb{R}^n$ so that A is diagonal in that basis.
- This involves the eigenvectors and eigenvalues of A.
- You can visualize the effect of matrix multiplication on vectors as consisting of two operations: i) a rotation and ii) a stretching or shrinking.
- Eigenvectors of A are those vectors that are not rotated, i.e. if $\neq 0$ is an eigenvector of A, then for some $\lambda \in \mathbb{R}$,

$$Ax = \lambda x.$$

Eigenvalues and eigenvectors of a matrix

- We may write this more compactly as

$$(A - \lambda I)x = 0,$$

  where I is the $n \times n$ identity matrix.
- But from basic linear algebra, we know that a homogenous linear equation can have a non-zero solution only if the matrix does not have full rank
- Or $\det(A - \lambda I) = 0$.
- The values of λ for which this determinant is zero are called the eigenvalues of A.

Eigenvalues and eigenvectors of a matrix

- The determinant of $(A - \lambda I)$ is called the characteristic polynomial of A so the eigenvalues are the roots of the characteristic polynomial.
- If A has n distinct eigenvalues $\lambda_1, ..., \lambda_n$, then it has also n linearly independent eigenvectors $\boldsymbol{v}_1, ..., \boldsymbol{v}_n$ so that

$$A\boldsymbol{v}_i = \lambda_i \boldsymbol{v}_i.$$

- Let' see an example on how to compute the eigenvalues and vectors.
- Let

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Eigenvalues and eigenvectors of a matrix

- Then
$$A - \lambda I = \begin{pmatrix} 1-\lambda & 1 \\ 1 & -\lambda \end{pmatrix},$$
and
$$\det(A - \lambda I) = \lambda^2 - \lambda - 1.$$
- We have $\det(A - \lambda I) = 0$ if
$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}.$$
- The corresponding eigenvectors are:
$$\boldsymbol{v}_1 = (\frac{1+\sqrt{5}}{2}, 1), \boldsymbol{v}_2 = (\frac{1-\sqrt{5}}{2}, 1).$$
- A useful thing to keep in mind about eigenvalues is that the sum of the eigenvalues equals the trace (i.e. the sum of diagonal elements) of the matrix and the product of the eigenvalues equals the determinant of the matrix.
- This is particularly useful for inference about the signs of eigenvalues.

Eigenvalues and eigenvectors of a matrix

- Since the characteristic polynomial may fail to have real roots, eigenvalues correspond to the case where the matrix does not have any directions that are not rotated.
- To see an easy example of such a matrix, consider the 90-degree rotation anticlockwise:

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

- The characteristic polynomial for this matrix is $\lambda^2 + 1$ which obviously does not have real roots.
- If the eigenvalues are complex numbers, the eigenvectors are also have complex coordinates.
- We do not have time in this course to pursue this, but it should be pointed out that the method outlined below for solving the difference equations extends also to the case with complex eigenvalues.

Eigenvectors, eigenvalues and difference equations

- I can express any $x \in \mathbb{R}^n$ given in the usual coordinate system in the coordinate system spanned by the eigenvectors by simple matrix multiplication.
- Let $P = [\boldsymbol{v}_1 \quad \boldsymbol{v}_2 \quad \ldots \boldsymbol{v}_n]$ be the matrix formed by the eigenvectors. Then for any vector y expressed in the coordinate system of the eigenvectors, we can translate it to the standard system by $x = Py$.
- Similarly any x in the standard system is $y = P^{-1}x$ in the system of the eigenvectors.
- Why is this helpful at all?

Eigenvectors, eigenvalues and difference equations

- Consider now how y_{t+1} (in the new coordinate system) depends on y_t.

$$y_{t+1} = P^{-1}x_{t+1} = P^{-1}Ax_t = P^{-1}APy_t.$$

- We want to show that $P^{-1}AP = \Lambda$, where Λ is the diagonal matrix of eigenvalues.
- But this is the same claim as (premultiply by P):

$$AP = P\Lambda.$$

- But this follows immediately from the fact that P consists of the eigenvectors of A. (Make sure you understand this by writing $P = [\boldsymbol{v}_1, ..., \boldsymbol{v}_n]$ and calculating the matrix product on both sides).

Eigenvectors, eigenvalues and difference equations

- Hence we have: $y_t = (y_{1,t}, ..., y_{n,t}) = (c_1\lambda_1^t, ..., c_n\lambda_n^t)$.
- Since $x_t = Py_t$, we have the general solution:

$$x_t = c_1\lambda_1^t \boldsymbol{v}_1 + ... + c_n\lambda_n^t \boldsymbol{v}_n.$$

- Note that $A^t = P\Lambda^t P^{-1}$.
- Therefore we could have also concluded that

$$x_t = P\Lambda^t P^{-1} x_0.$$

- The two methods give the same results since $Pc = x_0$ or $c = P^{-1}x_0$.

Eigenvectors, eigenvalues and difference equations

- Sometimes a matrix has a repeated eigenvalue. Consider for example

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

- Then the characteristic equation is $(1-\lambda)^2 = 0$ and the matrix has a single eigenvalue $\lambda = 1$ and therefore a single eigenvector $(1, 0)$.
- This matrix cannot be diagonalized in the procedure that we had above.
- Luckily enough all matrices can be expressed as

$$A = Q^{-1} B Q,$$

  where Q is a matrix of generalized eigenvalues and B is upper triangular (and even better in so called Jordan normal form).
- Since the powers of upper triangular matrices are easy to compute, the same procedure as before can be applied for solving the model. See the book for the details on this.

Properties of the solutions

- For all (homogenous) systems of linear difference equations, 0 is a steady state.
- If A has full rank, it is the only steady state. Does the system eventually converge to its steady state?
- Look at the general solution

$$x_t = c_1 \lambda_1^t v_1 + ... + c_n \lambda_n^t v_n.$$

Properties of the solutions

- If $|\lambda_i| < 1$ for all i, then $x_t \to 0$ for all c. We say that in this case, the origin is a globally stable steady state or a sink.
- If $|\lambda_i| > 1$ for all i, then the length of x_t grows without bound for all c. We say that the origin is unstable or a source.
- Finally if If $|\lambda_i| < 1$ for some i and If $|\lambda_i| > 1$ for some i, then the length of x_t grows without bound if $c_i \neq 0$ for some i with $|\lambda_i| > 1$. If $c \neq 0$ only for i with $|\lambda_i| < 1$, then x_t converges to the origin.
- In this last case, we say that origin is a saddle point for the system. If $\lambda_i = 1$ for some i, then origin is neither stable, unstable nor a saddle.

Linearizing non-linear systems

- For your future information, I note here that if x^* is a steady state of a nonlinear system, we can use Taylor's first order approximation to analyze the local behavior of the system around the steady state (you'll do this in macroeconomics a lot).
- Suppose that $x_{t+1} = f(x_t)$ and $x^* = f(x^*)$.
- Then we have

$$x_{t+1} = f(x_t) = f(x^*) + D_x f(x^*)(x_t - x^*) \text{ or}$$

$$x_{t+1} - x^* = D_x f(x^*)(x_t - x^*).$$

- But this is a linear system in the deviations from the steady state and we can apply the analysis from the linear case in the for small deviations.
- You can classify the steady states of nonlinear models locally as we just did for the linear system (but globally). Just look at the absolute values of the eigenvalues and compare to 1.

Examples: Markov model

- Consider the system
$$x_{t+1} = Px_t$$
for a stochastic matrix P, i.e. non-negative matrix whose elements in each column sum up to 1.
- You have already shown in Problem set 0 that $\lambda = 1$ is an eigenvalue for all Markov matrices.
- It can be shown that in the case with strictly positive entries, all other eigenvalues are less that one in absolute value. Therefore x_t converges in the long run to the eigenvector (whose coordinates are normalized to sum to 1) corresponding to eigenvalue 1.
- The second largest (in length) eigenvalue measures the speed of convergence to this eigenvector.

Examples: Markov model

- Let l_t denote the fraction of employed and u_t the fraction of unemployed population in period t. The following difference equation system describes the evolution of these fractions.

$$\begin{bmatrix} l_{t+1} \\ u_{t+1} \end{bmatrix} = \begin{bmatrix} 0.9 & 0.7 \\ 0.1 & 0.3 \end{bmatrix} \begin{bmatrix} l_t \\ u_t \end{bmatrix}$$

- We know from before that one of the eigenvalues for a stochastic matrix is 1 (note that the columns of he matrix both add up to 1).

Examples: Markov model

- Let's compute them anyhow.
- Characteristic polynomial:

$$(0.9-\lambda)(0.3-\lambda)-0.07=\lambda^2-1.2\lambda+0.2=(\lambda-1)(\lambda-0.2)$$

- (You could have found the other eigenvalue also by subtracting 1 from the trace of the matrix.)
- For λ_1, we get

$$\begin{aligned} -0.1v_1+0.7v_2&=0 \\ 0.1v_1-0.7v_2&=0 \end{aligned}$$

- We can pick any vector satisfying these.

$$\mathbf{v}_{\lambda_1}=\begin{bmatrix} 7 \\ 1 \end{bmatrix}$$

Examples: Markov model

- The eigenvector for λ_2 is solved from:

$$0.7v_1 + 0.7v_2 = 0$$
$$0.1v_1 + 0.1v_2 = 0$$

- For example: $\mathbf{v_2} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}$.
- We can write the general solution to this system of difference equations as

$$\begin{bmatrix} l_t \\ u_t \end{bmatrix} = c_1 1^t \mathbf{v_1} + c_2 (0.2)^t \mathbf{v_2}$$

Examples: Markov model

- Since $l_t + u_t = 1$ (since these are fractions, we must set $c_1 = \frac{1}{8}$.

$$\begin{bmatrix} l_\infty \\ u_\infty \end{bmatrix} = \begin{bmatrix} \frac{7}{8} \\ \frac{1}{8} \end{bmatrix}$$

- In words, in the long run, fraction $\frac{7}{8}$ are employed and fraction $\frac{1}{8}$ are unemployed.

Examples: Fibonacci sequence

- In yesterday's lecture, we talked about the Fibonacci sequence $x_0 = 0, x_1 = 1$

$$x_{t+2} = x_{t+1} + x_t \text{ for } t \geq 2.$$

- This can be written as the following system:

$$\begin{pmatrix} x_{t+2} \\ x_{t+1} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_{t+1} \\ x_t \end{pmatrix}.$$

- We saw already earlier that the eigenvalues are

$$\lambda_1 = \frac{1+\sqrt{5}}{2}, \lambda_2 = \frac{1-\sqrt{5}}{2}.$$

- The corresponding eigenvectors are:

$$\boldsymbol{v}_1 = (\frac{1+\sqrt{5}}{2}, 1), \boldsymbol{v}_2 = (\frac{1-\sqrt{5}}{2}, 1).$$

Examples: Fibonacci sequence

- Hence the general solution to the Fibonacci difference equation is

$$\begin{pmatrix} x_{t+1} \\ x_t \end{pmatrix} = c_1\left(\frac{1+\sqrt{5}}{2}\right)^t \boldsymbol{v}_1 + c_2\left(\frac{1-\sqrt{5}}{2}\right)^t \boldsymbol{v}_2.$$

- At $t = 0$, we have

$$\begin{pmatrix} x_{t+1} \\ x_t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \frac{1}{\sqrt{5}}(\boldsymbol{v}_1 - \boldsymbol{v}_2).$$

- Therefore the general solution is:

$$x_t = \frac{\left(\frac{1+\sqrt{5}}{2}\right)^t - \left(\frac{1-\sqrt{5}}{2}\right)^t}{\sqrt{5}}.$$