OSM Bootcamp

Lecture 3

John Stachurski

2018

QuantEcon

Vector Analysis: Preliminaries

Let $\mathbb{R}^n$ denote the set of all n vectors $x = (x_1, \ldots, x_n)$

- In matrix algebra, x defaults to column vector

The **Euclidean norm** $\| \cdot \|$ on $\mathbb{R}^n$ is defined by

$$\|x\| := \left(\sum_{i=1}^{n} x_i^2 \right)^{1/2}$$

Interpretation:

- $\|x\|$ represents the "length" of x
- $\|x - y\|$ represents distance between x and y

QuantEcon

Fact. For any $\alpha \in \mathbb{R}$ and any $x, y \in \mathbb{R}^n$, the following statements are true:

1. $\|x\| \geqslant 0$ and $\|x\| = 0$ if and only if $x = 0$ positivity and identification
2. $\|\alpha x\| = |\alpha| \|x\|$ linear homogeneity
3. $\|x + y\| \leqslant \|x\| + \|y\|$ (**triangle inequality**)
4. $|x'y| \leqslant \|x\| \|y\|$ (**Cauchy-Schwarz inequality**)

(Here $x'y$ is the **inner product** $\sum_{i=1}^n x_i y_i$)

The Set of Matrices $\mathcal{M}(n \times k)$

Let $\mathcal{M}(n \times k)$ be the set of $n \times k$ real matrices

Questions:

- When is matrix A "close" to matrix B?
- When does A_n converge to A?
- What does $\sum_{n=1}^{\infty} A_n$ mean?

To answer these questions, we introduce a norm on $\mathcal{M}(n \times k)$

The Spectral Norm

Given $A \in \mathcal{M}(n \times k)$, the **spectral norm** of A is

$$\|A\| := \sup \left\{ \frac{\|Ax\|}{\|x\|} \; : \; x \in \mathbb{R}^k, \; x \neq 0 \right\}$$

- LHS is the spectral norm of A
- RHS is ordinary Euclidean vector norms

We often just say the **norm** of A

QuantEcon

Properties of the Spectral Norm

Similar to Euclidean norms on vectors,

Fact. For all $A, B \in \mathcal{M}(n \times k)$,

1. $\|A\| \geqslant 0$ and $\|A\| = 0 \iff A = 0$
2. $\|\alpha A\| = |\alpha| \|A\|$ for any scalar α
3. $\|A + B\| \leqslant \|A\| + \|B\|$

Pick $x \neq 0$.

Then $\frac{\|Ax\|}{\|x\|} \leq \overbrace{\|A\|}^{\text{sup over all } x\text{'s}}$

$\|Ax\| \leq \|A\| \|x\|$

Ex. Show that

$$\|Ax\| \leqslant \|A\| \cdot \|x\| \qquad \forall\, x \in \mathbb{R}^k$$

conformable

Fact. If AB is well defined, then $\|AB\| \leqslant \|A\| \|B\|$

Proof: Let $A \in \mathcal{M}(n \times k)$, let $B \in \mathcal{M}(k \times j)$ and let $x \in \mathbb{R}^j$

We have

$$\|A\underbrace{Bx}_{\text{just a column vector}}\| \leqslant \|A\| \cdot \|Bx\| \leqslant \|A\| \cdot \|B\| \cdot \|x\|$$

$$\therefore \quad \frac{\|ABx\|}{\|x\|} \leqslant \|A\| \cdot \|B\|$$

holds for all x

=> holds for sup of x's

=> $\|AB\| \leq \|A\| \|B\|$

Called the **submultiplicative property**

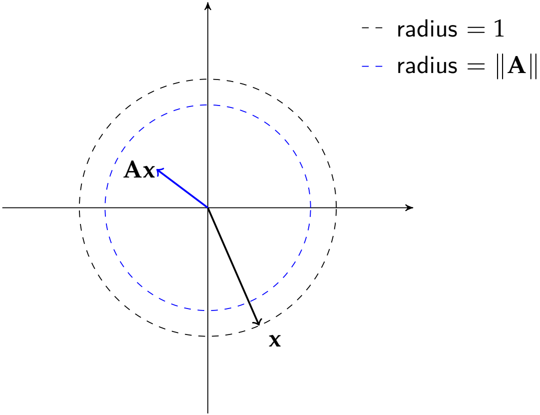
Implication: $\|A^j\| \leqslant \|A\|^j$ for any $j \in \mathbb{N}$ and $A \in \mathcal{M}(n \times n)$

square matrices

QuantEcon

If $\|A\| \leqslant 1$ then A is called **nonexpansive**

If $\|A\| < 1$ then A is called **contractive**



Distance, Convergence, etc.

Having a norm on matrices gives us a notion of distance:

$$d(A, B) = \|A - B\|$$

norm deviation

Example. If $\|A_j - A\| \to 0$ then we say that A_j **converges** to A

norm deviation goes to 0.

Similarly,

$$\sum_{j=1}^{\infty} A_j = B \quad \iff \quad \lim_{J \to \infty} \left\| \sum_{j=1}^{J} A_j - B \right\| = 0$$

QuantEcon

For $A \in \mathcal{M}(n \times n)$, the **spectral radius** is

$$r(A) := \max\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$$

Fact. For all $A \in \mathcal{M}(n \times n)$, we have

1. $\|A\| = \sqrt{r(A'A)}$
2. $\|A'\| = \|A\|$ and $r(A') = r(A)$

Fact. (**Gelfand's formula**) For all $A \in \mathcal{M}(n \times n)$, we have

$$\|A^k\|^{1/k} \to r(A) \quad \text{as} \quad k \to \infty$$

Ex. Use Gelfand's formula to show that

go back to definition of convergence (use ε and $N \in \mathbb{N}$)

$$r(A) < 1 \implies \|A^k\| \to 0$$

Neumann Series Lemma

Let $A \in \mathcal{M}(n \times n)$ and let I be the $n \times n$ identity

Fact. (**Neumann series lemma.**) If $r(A) < 1$, then $I - A$ is nonsingular and

$$(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$$

Example. If $r(A) < 1$, then $x = Ax + b$ has the unique solution

$$x^* = \sum_{j=0}^{\infty} A^j b$$

generalization of geometric series in scalar case:

$|a| < 1 \Rightarrow \sum_{t \geq 0} a^t = \frac{1}{1-a}$

$x = Ax + b$

$x(I - A) = b$

$x^* = (I - A)^{-1} b$

Proof of the NSL

Converge if underlying space complete

Ex. Show that $B_J := \sum_{j=0}^{J} A^j$ is Cauchy and hence $\sum_{j=0}^{\infty} A^j$ exists

Now observe that $(I - A) \sum_{j=0}^{\infty} A^j = I$, since

$$\left\| (I - A) \sum_{j=0}^{\infty} A^j - I \right\| = \left\| (I - A) \lim_{J \to \infty} \sum_{j=0}^{J} A^j - I \right\|$$

$$= \lim_{J \to \infty} \left\| (I - A) \sum_{j=0}^{J} A^j - I \right\|$$

$$= \lim_{J \to \infty} \left\| A^{J+1} \right\| = 0$$

QuantEcon

Linear Vector-Valued Systems

Let $A \in \mathcal{M}(n \times n)$ and consider the dynamic model

$$x_{t+1} = Ax_t + b, \quad x_0 \text{ given}$$

Example. Next period inflation and output depend on current inflation and output via certain laws of motion

As a **dynamical system**,

- $\mathbb{X} = \mathbb{R}^n$
- $g(x) = Ax + b$

QuantEcon

As before, a steady state is a vector x^* such that $x^* = g(x^*)$

That is,

$$x^* = Ax^* + b$$

then has a unique solution

Fact. If $r(A) < 1$, then $(\mathbb{X}, g)$ is **globally stable**, with unique steady state

$$x^* = \sum_{j=0}^{\infty} A^j b$$

Existence and uniqueness follows from the Neumann Series Lemma

$x_2 = Ax_1 + b$

$= A(Ax_0 + b) + b$

$= A^2 x_0 + Ab + b$

$\vdots$

How about stability? Iteration gives

$$x_t = A^t x_0 + A^{t-1} b + \cdots + b$$

Hence, for any x_0, y_0 in $\mathbb{R}^n$, we have

effect of initial conditions

$$\begin{aligned} \|x_t - y_t\| &= \left\|A^t x_0 - A^t y_0\right\| \\ &= \left\|A^t (x_0 - y_0)\right\| \\ &\leqslant \|A^t\| \cdot \|x_0 - y_0\| \end{aligned}$$

sub multiplicative

Using $\underline{r(A) < 1}$ and setting $\underline{y_0 = x^*}$ gives $x_t \to x^*$

key condition

every point converging to x^*

QuantEcon

Linear Vector Systems with Noise

Next consider

$n \times j$ $j \times 1$ => $C\xi_{t+1}$ is an $n \times 1$ vector

- $x_{t+1} = Ax_t + b + C\xi_{t+1}$ with x_0 given
- $\{\xi_t\}$ is IID and satisfies

$$\mathbb{E}[\xi_{t+1}] = 0 \quad \text{and} \quad \mathbb{E}[\xi_{t+1}\xi_{t+1}'] = I$$

elementwise expectation

no correlation

What is the time path of the first two moments

- $\mu_t := \mathbb{E}[x_t]$
- $\Sigma_t := \text{var}[x_t] := \mathbb{E}[(x_t - \mu_t)(x_t - \mu_t)']$

QuantEcon

Dynamics of the Mean

First, regarding μ_t, take expectations over

$$x_{t+1} = Ax_t + b + C\xi_{t+1}$$

to get

$$\mu_{t+1} = A\mu_t + b$$

deterministic linear system.

Fact. If $r(A) < 1$, then $\{\mu_t\}$ converges to the unique fixed point

$$\mu^* = \sum_{i=0}^{\infty} A^i b$$

regardless of μ_0

QuantEcon

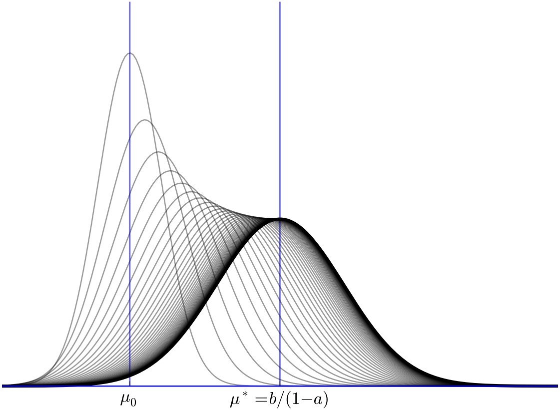


Figure: Convergence of μ_t to μ^* in the scalar model

Dynamics of the Variance

Consider again

$$x_{t+1} = Ax_t + b + C\xi_{t+1}$$

We want a similar law of motion for $\Sigma_t := \mathrm{var}[x_t]$

We will use the fact that $\mathbb{E}\left[x_t \xi'_{t+1}\right] = 0$

This follows from the assumptions above

x_t is a function of initial condition and all past shocks

$\Rightarrow$ independent of ξ_{t+1}

By definition,

$$\begin{aligned}\operatorname{var}[x_{t+1}] &= \mathbb{E}\left[(x_{t+1} - \mu_{t+1})(x_{t+1} - \mu_{t+1})'\right] \\ &= \mathbb{E}\left[(A(x_t - \mu_t) + C\xi_{t+1})(A(x_t - \mu_t) + C\xi_{t+1})'\right]\end{aligned}$$

The right hand side is equal to

$$\mathbb{E}\left[A(x_t - \mu_t)(x_t - \mu_t)'A'\right] + \mathbb{E}\left[A(x_t - \mu_t)\xi_{t+1}'C'\right] \\ + \mathbb{E}\left[C\xi_{t+1}(x_t - \mu_t)'A'\right] + \mathbb{E}\left[C\xi_{t+1}\xi_{t+1}'C'\right]$$

Some further manipulations (check) lead to

Difference equation in matrix space.

$$\Sigma_{t+1} = A\Sigma_t A' + CC'$$

QuantEcon

To repeat

$$\Sigma_{t+1} = g(\Sigma_t) \quad \text{where} \quad g(\Sigma) = A\Sigma A' + CC'$$

Variance is a trajectory of the dynamical system $(\mathcal{M}(n \times n), g)$

A steady state of this system is a Σ satisfying

$$\Sigma = A\Sigma A' + CC'$$

Fact. If $r(A) < 1$, then $(\mathcal{M}(n \times n), g)$ is **globally stable**

QuantEcon

More generally, consider the **discrete Lyapunov equation**

$$\Sigma = A \Sigma A' + M$$

- all matrices are in $\mathcal{M}(n \times n)$ and Σ is the unknown

Given A and M, let ℓ be the **Lyapunov operator**

$$\ell \colon \mathcal{M}(n \times n) \ni \Sigma \mapsto A \Sigma A' + M \in \mathcal{M}(n \times n)$$

Fact. If $r(A) < 1$, then $(\mathcal{M}(n \times n), \ell)$ is globally stable contraction mapping argument

Proof: Suffices to show that ℓ^k is a Banach contraction on $(\mathcal{M}(n \times n), \| \cdot \|)$ for some $k \in \mathbb{N}$

From the definition,

$$\ell^k(\Sigma) = A^k \Sigma (A^k)' + A^{k-1} M (A^{k-1})' + \cdots + M$$

Hence, for any Σ, Λ in $\mathcal{M}(n \times n)$, we have

$$\begin{aligned} \|\ell^k(\Sigma) - \ell^k(\Lambda)\| &= \left\| A^k \Sigma (A^k)' - A^k \Lambda (A^k)' \right\| \\ &= \left\| A^k (\Sigma - \Lambda)(A^k)' \right\| \\ &\leqslant \|A^k\| \cdot \|\Sigma - \Lambda\| \cdot \|(A^k)'\| \end{aligned}$$

QuantEcon

Transposes don't change norms, so $\|(A^k)'\| = \|A^k\|$ and hence

$$\|\ell^k(\Sigma) - \ell^k(\Lambda)\| \leqslant \|A^k\|^2 \|\Sigma - \Lambda\|$$

Since $r(A) < 1$, we can find $k \in \mathbb{N}$, $\lambda < 1$ such that

$$\|\ell^k(\Sigma) - \ell^k(\Lambda)\| \leqslant \lambda \|\Sigma - \Lambda\| \quad \text{for all } \Sigma, \Lambda \in \mathcal{M}(n \times m)$$

Now apply Banach contraction mapping theorem

Note: Gives an algorithm for computing Σ^*

(Not always the best one)

Stochastic Processes: Key Ideas

Quizz: Whose favorite saying is this? Sargent

An economic model is a probability distribution on a sequence space

But what's a probability distribution on a sequence space?

Let's break this down and try to understand...

X_t a random variable

Consider a **economic model** of the form

$$X_{t+1} = F(X_t, \xi_{t+1}), \qquad \text{where } \{\xi_t\} \overset{\text{IID}}{\sim} \phi$$

Objects such as F and ϕ are determined by theory + estimation + calibration

Here

- X_t is called the **state variable**
- It takes values in **state space** $\mathbb{X}$
- ξ_t is called the **shock** or **innovation**

QuantEcon

An economic model is a probability distribution on a sequence space

The "sequence space" is

$$\times_{t=0}^{\infty} \mathbb{X} := \mathbb{X} \times \mathbb{X} \times \mathbb{X} \times \cdots$$

A typical element is

$$(x_0, x_1, x_2, \ldots) \quad \text{where each } x_t \in \mathbb{X}$$

This is the set of all possible values for the time series

$$\mathbf{X} := (X_0, X_1, X_2, \ldots)$$

QuantEcon

The "probability distribution" on this sequence space is a map $\mathbb{P}_x$, where

$$\mathbb{P}_x(B) = \text{Prob}\{(X_0, X_1, X_2, \ldots) \in B\}$$

Here

- B is some "event" in the sequence space $\times_{t=0}^{\infty} \mathbb{X}$
- Prob means "probability of"

The subscript x in $\mathbb{P}_x$ means that we are conditioning on $X_0 = x$

An economic model is a probability distribution on a sequence space

Our economic model is $X_{t+1} = F(X_t, \xi_{t+1})$ with $\{\xi_t\} \overset{\text{IID}}{\sim} \phi$

The model determines the probability distribution $\mathbb{P}_x$ via

$$\mathbb{P}_x(B) = \text{Prob}\,\{(x, F(x, \xi_1), F(F(x, \xi_1), \xi_2), \ldots) \in B\}$$

This is the probability of the shock path

$$\{(z_1, z_2, \ldots) \mid (x, F(x, z_1), F(F(x, z_1), z_2), \ldots) \in B\}$$

according to the distribution $\times_{t=1}^{\infty} \phi$

QuantEcon

The distribution $\mathbb{P}_x$ tells us probabilities for the **whole path** $\{X_t\}$

It is the **joint distribution** of the sequence $\{X_t\}$

In theory, $\mathbb{P}_x$ can be used to answer any question along the lines

> "What's the probability that event B happens when $\{X_t\}$ is realized?"

Example. What's the probability that inflation falls each quarter for the next two years?

QuantEcon

Example. Inventory dynamics

- See Wk2_Dynamics/inventory_dynamics.ipynb

Example. Samuelson multiplier–accelerator with stochastic govt spending

- See Wk2_Dynamics/accelerator.ipynb

Marginal Distributions

Some events concern only one point in time

Let

$$\psi_t(B) := \mathbb{P}_x\{X_t \in B\} \quad \text{where} \quad B \subset \mathbb{X}$$

This object ψ_t is called the **marginal distribution** of X_t

Intuitively, $\psi_t(B)$ is

- the frequency of X_t landing in B if we run the system many times
- the fraction of "particles" that lie in B if many independent particles are generated by the model

Applications: See the discussion of marginal distributions in

- Wk2_Dynamics/inventory_dynamics.ipynb
- Wk2_Dynamics/accelerator.ipynb