

# Shenzhen Winter Camp

## Lecture 3

John Stachurski

2018

# Vector Analysis: Preliminaries

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

- 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  $y$

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

1.  $\|x\| \geq 0$  and  $\|x\| = 0$  if and only if  $x = 0$
2.  $\|\alpha x\| = |\alpha| \|x\|$
3.  $\|x + y\| \leq \|x\| + \|y\|$  (**triangle inequality**)
4.  $|x'y| \leq \|x\| \|y\|$  (**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  $A$

# Properties of the Spectral Norm

Similar to Euclidean norms on vectors,

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

1.  $\|A\| \geq 0$  and  $\|A\| = 0 \iff A = 0$
2.  $\|\alpha A\| = |\alpha| \|A\|$  for any scalar  $\alpha$
3.  $\|A + B\| \leq \|A\| + \|B\|$

**Ex.** Show that

$$\|Ax\| \leq \|A\| \cdot \|x\| \quad \forall x \in \mathbb{R}^k$$

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

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

$$\|ABx\| \leq \|A\| \cdot \|Bx\| \leq \|A\| \cdot \|B\| \cdot \|x\|$$

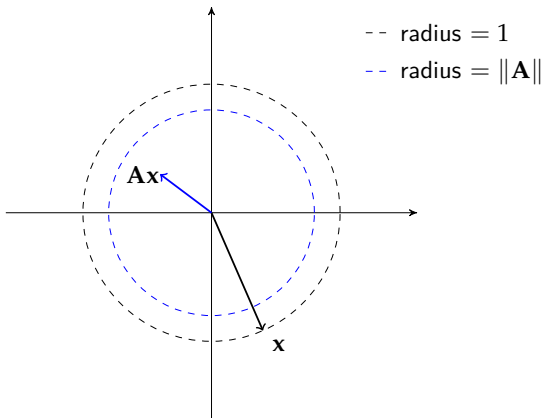
$$\therefore \frac{\|ABx\|}{\|x\|} \leq \|A\| \cdot \|B\|$$

Called the **submultiplicative property**

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

If  $\|A\| \leq 1$  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:

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

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

Similarly,

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

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** states that, for all  $A \in \mathcal{M}(n \times n)$ ,

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

**Ex.** Use Gelfand's formula to show that

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

# 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

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

## Proof of the NSL

**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

$$\begin{aligned} \left\| (I - A) \sum_{j=0}^{\infty} A^j - I \right\| &= \left\| (I - A) \lim_{J \rightarrow \infty} \sum_{j=0}^J A^j - I \right\| \\ &= \lim_{J \rightarrow \infty} \left\| (I - A) \sum_{j=0}^J A^j - I \right\| \\ &= \lim_{J \rightarrow \infty} \|A^{J+1}\| = 0 \end{aligned}$$

# 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**,

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

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

That is,

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

**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

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

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

Using  $r(A) < 1$  and setting  $y_0 = x^*$  gives  $x_t \rightarrow x^*$

# Linear Vector Systems with Noise

Next consider

- $x_{t+1} = Ax_t + b + C\tilde{\zeta}_{t+1}$  with  $x_0$  given
- $\{\tilde{\zeta}_t\}$  is IID and satisfies

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

What is the time path of the first two moments

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



# Dynamics of the Mean

First, regarding  $\mu_t$ , take expectations over

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

to get

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

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

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

regardless of  $\mu_0$

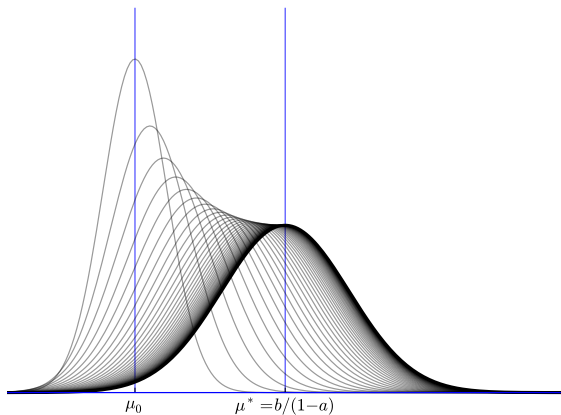


Figure: Convergence of  $\mu_t$  to  $\mu^*$  in the scalar model

# Dynamics of the Variance

Consider again

$$x_{t+1} = Ax_t + b + C\tilde{\zeta}_{t+1}$$

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

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

This follows from the assumptions above

By definition,

$$\begin{aligned}\text{var}[x_{t+1}] &= \mathbb{E}[(x_{t+1} - \mu_{t+1})(x_{t+1} - \mu_{t+1})'] \\ &= \mathbb{E}[(A(x_t - \mu_t) + C\tilde{\xi}_{t+1})(A(x_t - \mu_t) + C\tilde{\xi}_{t+1})']\end{aligned}$$

The right hand side is equal to

$$\begin{aligned}\mathbb{E}[A(x_t - \mu_t)(x_t - \mu_t)'A'] &+ \mathbb{E}[A(x_t - \mu_t)\tilde{\xi}_{t+1}'C'] \\ &+ \mathbb{E}[C\tilde{\xi}_{t+1}(x_t - \mu_t)'A'] + \mathbb{E}[C\tilde{\xi}_{t+1}\tilde{\xi}_{t+1}'C']\end{aligned}$$

Some further manipulations (check) lead to

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

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  $\Sigma$  satisfying

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

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

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

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

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

Given  $A$  and  $M$ , let  $\ell$  be the **Lyapunov operator**

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

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

Proof: Suffices to show that  $\ell^k$  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  $\Sigma, \Lambda$  in  $\mathcal{M}(n \times n)$ , we have

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

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

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

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

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

Note: Gives an algorithm for computing  $\Sigma^*$

(Not always the best one)



# Stochastic Processes: Key Ideas

Quiz: Whose favorite saying is this?

**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...

Consider a **economic model** of the form

$$X_{t+1} = F(X_t, \xi_{t+1}), \quad \text{where } \{\xi_t\} \stackrel{\text{iid}}{\sim} \phi$$

Objects such as  $F$  and  $\phi$  are determined by theory + estimation + calibration

Here

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

## An economic model is a probability distribution on a sequence space

The “sequence space” is

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

A typical element is

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

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

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

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

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

Here

- $B$  is some “event” in the sequence space  $\times_{t=0}^{\infty} \mathbb{X}$
- $\text{Prob}$  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  $\{\xi_t\} \stackrel{\text{iid}}{\sim} \phi$

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

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

This is the probability of the shock path

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

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

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?

### Example. Inventory dynamics

- See [John/inventory\\_dynamics.ipynb](#)

### Example. Samuelson multiplier–accelerator with stochastic govt spending

- See [John/accelerator.ipynb](#)

# Marginal Distributions

Some events concern only one point in time

Let

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

This object  $\psi_t$  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

- `John/inventory_dynamics.ipynb`
- `John/accelerator.ipynb`

Recall our model

$$X_{t+1} = F(X_t, \xi_{t+1}), \quad \text{where } \{\xi_t\} \stackrel{\text{iid}}{\sim} \phi$$

This model is **first order Markov**, which means that the marginal distribution  $\psi_{t+1}$  is fully determined by the model and  $\psi_t$

In particular,

$$\psi_{t+1} \stackrel{\mathcal{D}}{=} F(X, \xi) \quad \text{when} \quad (X, \xi) \stackrel{\mathcal{D}}{=} \psi_t \times \phi$$

**Example.** A **linear Gaussian AR(1) process** has the form

$$X_{t+1} = \rho X_t + b + \sigma \zeta_{t+1}, \quad \text{where } \{\zeta_t\} \stackrel{\text{iid}}{\sim} N(0, 1)$$

If  $\psi_t = N(\mu, s^2)$ , then

$$\psi_{t+1} = ?$$