ECON2125/4021/8013

Lecture 24

John Stachurski

Semester 1, 2015

Today's Lecture

An application of stochastic dynamics: Asset pricing

Moving average representations

Dynamics of stochastic systems

- Dynamics of moments
- Convergence of moments
- Dynamics of distributions, etc.

We start with some preliminaries

Preliminary 1: Expectation and Trace

In our application, we'll make use of the following result

Fact. If $\mathbf{w}$ is a random vector with $\mathbb{E}[\mathbf{w}\mathbf{w}'] = \mathbf{I}$ and $\mathbf{Q}$ is any conformable matrix, then

$$\mathbb{E}[\mathbf{w}'\mathbf{Q}\mathbf{w}] = \operatorname{trace}(\mathbf{Q})$$

Proof: Let q_{ij} be the i, j-th element of $\mathbf{Q}$

Note that

$$\mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_N \end{pmatrix} \quad \Longrightarrow \quad \mathbf{w}\mathbf{w}' = \begin{pmatrix} w_1 w_1 & \cdots & w_1 w_N \\ & \vdots & \\ w_N w_1 & \cdots & w_N w_N \end{pmatrix}$$

Hence

$$\mathbb{E}\left[\mathbf{w}\mathbf{w}'\right] = \mathbf{I} \quad \Longrightarrow \quad \mathbb{E}\left[w_i w_j\right] = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Now recall that

$$\mathbf{w}'\mathbf{Q}\mathbf{w} = \sum_{j=1}^{N}\sum_{i=1}^{N} q_{ij} w_i w_j$$

So, by linearity of expectations,

$$\mathbb{E}\left[\mathbf{w}'\mathbf{Q}\mathbf{w}\right] = \sum_{j=1}^{N}\sum_{i=1}^{N} q_{ij}\mathbb{E}\left[w_i w_j\right]$$

The result now follows

Preliminary 2: Lyapunov Equations

So far we've considered equations that have vectors as solutions

Sometimes we face equations that have matrices as solutions

An example is the **discrete Lyapunov equation**

$$\mathbf{P} = \mathbf{A}'\mathbf{P}\mathbf{A} + \mathbf{Q} \tag{1}$$

Here

- all matrices are $N \times N$
- $\mathbf{A}$ and $\mathbf{Q}$ are given
- $\mathbf{P}$ is the unknown

The question is, when does there exist a unique $\mathbf{P}$ that solves (1)?

Fact. Let $\mathbf{Q}$ and $\mathbf{A}$ be $N \times N$. If, in addition,

1. $\mathbf{Q}$ is symmetric
2. $\rho(\mathbf{A}) < 1$

then there exists a unique $\mathbf{P}$ that solves $\mathbf{P} = \mathbf{A}'\mathbf{P}\mathbf{A} + \mathbf{Q}$

If $\mathbf{Q}$ is positive definite, then so is the solution $\mathbf{P}$

Sketch of proof:

We studied the Banach contraction mapping theorem for vectors

Similar ideas carry through to matrices

Assumption $\rho(\mathbf{A}) < 1$ is used to obtain the contraction property

Application: Asset Pricing

Let's consider the problem of pricing an asset

- a house
- a firm
- a share in a firm, etc.

From a modeling perspective, an **asset** is a claim to a stream of payments, such as dividends

- a random sequence $\{d_t\}_{t=1}^{\infty}$

Our question:

What to pay at t for a claim to the dividend stream $d_{t+1}, d_{t+2}, \ldots$?

To answer this we need to take a stand on how dividends evolve

Let's assume that

1. $d_t = \mathbf{x}_t' \mathbf{D} \mathbf{x}_t$ for some positive definite $\mathbf{D}$
2. $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{C}\mathbf{w}_{t+1}$ for all t

Assumptions as before, including

- $\{\mathbf{w}_t\}$ is an MDS
- $\mathbb{E}_t[\mathbf{w}_{t+1}\mathbf{w}_{t+1}'] = \mathbf{I}$ for all t

Here $\mathbf{x}_t$ is a vector of random factors believed to affect dividends

- Investment growth in China? Price of oil?

Notice the functional form in

$$d_t = \mathbf{x}_t' \mathbf{D} \mathbf{x}_t$$

Why are we assuming that d_t is quadratic in the factors $\mathbf{x}_t$?

The short answer is simplicity

- we can still hope to find prices using algebra

So, if we want simplicity, why not assume that d_t is linear in $\mathbf{x}_t$?

This is simpler but too unrealistic

- e.g., can get negative dividends

Quadratic (with pos. definite $\mathbf{D}$) balances simplicity and realism

Risk Neutral Pricing

We are going to price the asset with "risk neutral" pricing

In our setting, this says that the price should satisfy

$$p_t = \beta \mathbb{E}_t[d_{t+1}] + \beta \mathbb{E}_t p_{t+1}$$

for all t, where

- p_k is price at time k
- $\beta \in (0, 1)$ discounts next period values to current
- $\mathbb{E}_t$ is the expectation given time t information

Note: This is a recursive representation of prices

We still have to work out p_t in terms of primitives

Predicting Quadratic Functions

One thing we need to do is predict future dividends

We want to predict from current information, so let's use $\mathbb{E}_t$

Let's start by predicting $\mathbf{x}_{t+1}'\mathbf{H}\mathbf{x}_{t+1}$ for arbitrary $\mathbf{H}$

We have

$$\mathbb{E}_t[\mathbf{x}_{t+1}'\mathbf{H}\mathbf{x}_{t+1}] = \mathbb{E}_t[(\mathbf{A}\mathbf{x}_t + \mathbf{C}\mathbf{w}_{t+1})'\mathbf{H}(\mathbf{A}\mathbf{x}_t + \mathbf{C}\mathbf{w}_{t+1})]$$

Ex. Expand the right hand side out to get

$$\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{A}\mathbf{x}_t] + 2\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}] + \mathbb{E}_t[\mathbf{w}_{t+1}'\mathbf{C}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}]$$

Hint: A scalar valued expression is equal to its transpose

So consider the expression

$$\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{A}\mathbf{x}_t] + 2\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}] + \mathbb{E}_t[\mathbf{w}_{t+1}'\mathbf{C}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}]$$

Regarding the first term, since $\mathbf{x}_t$ is known at t we have

$$\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{A}\mathbf{x}_t] = \mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{A}\mathbf{x}_t$$

Regarding the second, since $\{\mathbf{w}_t\}$ is an MDS,

$$2\mathbb{E}_t[\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}] = 2\mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{C}\,\mathbb{E}_t[\mathbf{w}_{t+1}] = 0$$

Regarding the third, we can use our result from the start of the lecture to get

$$\mathbb{E}_t[\mathbf{w}_{t+1}'\mathbf{C}'\mathbf{H}\mathbf{C}\mathbf{w}_{t+1}] = \mathrm{trace}(\mathbf{C}'\mathbf{H}\mathbf{C})$$

Predicting Dividends

Combining these results gives our final expression

$$\mathbb{E}_t[\mathbf{x}_{t+1}'\mathbf{H}\mathbf{x}_{t+1}] = \mathbf{x}_t'\mathbf{A}'\mathbf{H}\mathbf{A}\mathbf{x}_t + \mathrm{trace}(\mathbf{C}'\mathbf{H}\mathbf{C})$$

Applying this to prediction of dividends gives

$$\mathbb{E}_t[d_{t+1}] = \mathbf{x}_t'\mathbf{A}'\mathbf{D}\mathbf{A}\mathbf{x}_t + \mathrm{trace}(\mathbf{C}'\mathbf{D}\mathbf{C})$$

Comments

- Our time t prediction of d_{t+1} is a function of $\mathbf{x}_t$
- The same can be shown for predictions of any d_{t+j}

Prices as Functions of the State

We've seen that all information useful for predicting future dividends is contained in $\mathbf{x}_t$

This leads us to conjecture that p_t should be a function of $\mathbf{x}_t$

- Prices are functions of data useful for predicting dividends

We're going to make another leap and guess that prices are a quadratic in $\mathbf{x}_t$

In particular, we guess that the solution p_t takes the form

$$p_t = \mathbf{x}_t' \mathbf{V} \mathbf{x}_t + \delta$$

for some positive definite $\mathbf{V}$ and scalar δ

The plan: See if there exist $\mathbf{V}$ and δ such that

$$p_t = \mathbf{x}_t' \mathbf{V} \mathbf{x}_t + \delta \tag{2}$$

satisfies the risk neutral pricing equation

$$p_t = \beta \mathbb{E}_t [d_{t+1}] + \beta \mathbb{E}_t p_{t+1}$$

Substituting (2) into both sides gives

$$\mathbf{x}_t' \mathbf{V} \mathbf{x}_t + \delta = \beta \mathbb{E}_t [\mathbf{x}_{t+1}' \mathbf{D} \mathbf{x}_{t+1}] + \beta \mathbb{E}_t [\mathbf{x}_{t+1}' \mathbf{V} \mathbf{x}_{t+1} + \delta]$$

Ex. Show from our results on predicting quadratics that gives

$$\mathbf{x}_t' \mathbf{V} \mathbf{x}_t + \delta = \beta \mathbf{x}_t' \mathbf{A}' (\mathbf{D} + \mathbf{V}) \mathbf{A} \mathbf{x}_t + \beta \operatorname{trace}(\mathbf{C}' (\mathbf{D} + \mathbf{V}) \mathbf{C}) + \beta \delta$$

So, we seek a pair $\mathbf{V}$, δ that solves

$$\mathbf{x}_t'\mathbf{V}\mathbf{x}_t + \delta = \beta \mathbf{x}_t'\mathbf{A}'(\mathbf{D}+\mathbf{V})\mathbf{A}\mathbf{x}_t + \beta \operatorname{trace}(\mathbf{C}'(\mathbf{D}+\mathbf{V})\mathbf{C}) + \beta\delta$$

for any $\mathbf{x}_t$

<u>Suppose</u> that there exists an $N \times N$ matrix $\mathbf{V}^*$ such that

$$\mathbf{V}^* = \beta \mathbf{A}'(\mathbf{D}+\mathbf{V}^*)\mathbf{A}$$

<u>Claim:</u> If this is true and we define δ^* as

$$\delta^* := \frac{\beta}{1-\beta} \operatorname{trace}(\mathbf{C}'(\mathbf{D}+\mathbf{V}^*)\mathbf{C})$$

then the pair $\mathbf{V}^*, \delta^*$ solves the above equation for any $\mathbf{x}_t$

Proof: By hypothesis, $\mathbf{V}^* = \beta \mathbf{A}'(\mathbf{D} + \mathbf{V}^*)\mathbf{A}$

$$\therefore \quad \mathbf{x}_t' \mathbf{V}^* = \beta \mathbf{x}_t' \mathbf{A}'(\mathbf{D} + \mathbf{V}^*)\mathbf{A}$$

$$\therefore \quad \mathbf{x}_t' \mathbf{V}^* \mathbf{x}_t = \beta \mathbf{x}_t' \mathbf{A}'(\mathbf{D} + \mathbf{V}^*)\mathbf{A}\mathbf{x}_t$$

$$\therefore \quad \mathbf{x}_t' \mathbf{V}^* \mathbf{x}_t + \delta^* = \beta \mathbf{x}_t' \mathbf{A}'(\mathbf{D} + \mathbf{V}^*)\mathbf{A}\mathbf{x}_t + \delta^*$$

By definition,

$$\delta^* := \frac{\beta}{1-\beta} \operatorname{trace}(\mathbf{C}'(\mathbf{D} + \mathbf{V}^*)\mathbf{C})$$

Ex. Complete the proof

Hence the problem comes down to finding a $\mathbf{V}$ that solves

$$\mathbf{V} = \beta \mathbf{A}'(\mathbf{D} + \mathbf{V})\mathbf{A} \tag{3}$$

Claim: A unique solution to (3) exists whenever $\rho(\sqrt{\beta}\mathbf{A}) < 1$

Proof: Letting $\mathbf{Q} := \beta \mathbf{A}'\mathbf{D}\mathbf{A}$ and $\mathbf{\Lambda} := \sqrt{\beta}\mathbf{A}$, we can express (3) as

$$\mathbf{V} = \mathbf{\Lambda}'\mathbf{V}\mathbf{\Lambda} + \mathbf{Q}$$

- A discrete Lyapunov equation in $\mathbf{V}$
- Since $\mathbf{D}$ is symmetric (being positive definite), so is $\mathbf{Q}$

Since $\rho(\mathbf{\Lambda}) < 1$, a unique solution $\mathbf{V}$ exists

Ex. Show that $\mathbf{V}$ is positive definite whenever $\mathbf{A}$ is nonsingular

Asset Pricing Summary

We started with the risk neutral asset pricing equation

$$p_t = \beta \mathbb{E}_t[d_{t+1}] + \beta \mathbb{E}_t p_{t+1}$$

with

$$d_t = \mathbf{x}_t' \mathbf{D} \mathbf{x}_t, \qquad \mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{C}\mathbf{w}_{t+1}$$

We have shown that

$$\rho(\sqrt{\beta}\mathbf{A}) < 1 \implies \mathbf{V} = \beta \mathbf{A}'(\mathbf{D} + \mathbf{V})\mathbf{A} \text{ has a unique solution}$$

From the solution $\mathbf{V}^*$ and an associated constant δ^* we get a solution

$$p_t^* := \mathbf{x}_t' \mathbf{V}^* \mathbf{x}_t + \delta^*$$

Moving Average Representations

Now let's return to the general case where

- $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1}$
- $\mathbf{w}_t$ is a MDS with $\mathbb{E}_t[\mathbf{w}_{t+1}\mathbf{w}'_{t+1}] = \mathbf{I}$
- $\mathbf{x}_0$ is a constant

In the deterministic case we expressed $\mathbf{x}_t$ in terms of $\mathbf{x}_0$

Here we can express $\mathbf{x}_t$ in terms of $\mathbf{x}_0$ and $\mathbf{w}_1, \ldots, \mathbf{w}_t$

Letting $\mathbf{v}_t := \mathbf{b} + \mathbf{C}\mathbf{w}_t$, we have

$$
\begin{aligned}
\mathbf{x}_t &= \mathbf{A}\mathbf{x}_{t-1} + \mathbf{v}_t \\
&= \mathbf{A}(\mathbf{A}\mathbf{x}_{t-2} + \mathbf{v}_{t-1}) + \mathbf{v}_t \\
&= \mathbf{A}^2\mathbf{x}_{t-2} + \mathbf{A}\mathbf{v}_{t-1} + \mathbf{v}_t \\
&= \mathbf{A}^2(\mathbf{A}\mathbf{x}_{t-3} + \mathbf{v}_{t-2}) + \mathbf{A}\mathbf{v}_{t-1} + \mathbf{v}_t \\
&= \mathbf{A}^3\mathbf{x}_{t-3} + \mathbf{A}^2\mathbf{v}_{t-2} + \mathbf{A}\mathbf{v}_{t-1} + \mathbf{v}_t \\
&= \cdots
\end{aligned}
$$

More generally,

$$\mathbf{x}_t = \mathbf{A}^j \mathbf{x}_{t-j} + \mathbf{A}^{j-1}\mathbf{v}_{t-(j-1)} + \mathbf{A}^{j-2}\mathbf{v}_{t-(j-2)} + \cdots + \mathbf{A}\mathbf{v}_{t-1} + \mathbf{v}_t$$

Setting $j = t$,

$$\begin{aligned}\mathbf{x}_t &= \mathbf{A}^t \mathbf{x}_0 + \mathbf{A}^{t-1}\mathbf{v}_1 + \mathbf{A}^{t-2}\mathbf{v}_2 + \cdots + \mathbf{A}\mathbf{v}_{t-1} + \mathbf{v}_t \\ &= \mathbf{A}^t \mathbf{x}_0 + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{v}_{t-i}\end{aligned}$$

Making the substitution $\mathbf{v}_{t-i} = \mathbf{b} + \mathbf{C}\mathbf{w}_{t-i}$, we get

$$\mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{b} + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{C}\mathbf{w}_{t-i}$$

The expression

$$\mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{b} + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{C} \mathbf{w}_{t-i} \tag{4}$$

is called the **moving average** or **MA** representation of $\mathbf{x}_t$

Example. Consider the scalar case $x_t = a x_{t-1} + w_t$ with $|a| < 1$

The MA representation is

$$x_t = a^t x_0 + a^{t-1} w_1 + a^{t-2} w_2 + \cdots + a w_{t-1} + w_t$$

Since $|a| < 1$, earlier shocks (e.g., w_1) have less influence than later ones (e.g., w_t)

- Similar story in (4) when $\|\mathbf{A}\| < 1$

Dynamics of Moments

Because of the shocks in

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1}$$

we don't know exactly what will happen to $\{\mathbf{x}_t\}$

- Perturbed by shocks at each t

But we can work out the time path of the first two moments

- $\boldsymbol{\mu}_t := \mathbb{E}[\mathbf{x}_t]$
- $\boldsymbol{\Sigma}_t := \mathrm{var}[\mathbf{x}_t] := \mathbb{E}[(\mathbf{x}_t - \boldsymbol{\mu}_t)(\mathbf{x}_t - \boldsymbol{\mu}_t)']$

These sequences are nonrandom

Dynamics of the Mean

First, regarding $\boldsymbol{\mu}_t$, take expectations over

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1}$$

to get

$$\mathbb{E}\left[\mathbf{x}_{t+1}\right] = \mathbb{E}\left[\mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1}\right] = \mathbf{A}\mathbb{E}\left[\mathbf{x}_t\right] + \mathbf{b}$$

In other words,

$$\boldsymbol{\mu}_{t+1} = \mathbf{A}\boldsymbol{\mu}_t + \mathbf{b}$$

This linear difference equation tells us how $\boldsymbol{\mu}_t$ evolves

Dynamics of the Variance

We want a similar law of motion for $\mathbf{\Sigma}_t := \mathrm{var}[\mathbf{x}_t]$

In finding it we'll use the following fact

Fact. Under our assumptions, $\mathbb{E}\left[\mathbf{x}_t \mathbf{w}'_{t+1}\right] = \mathbf{0}$ for all t

Proof: From the law of iterated expectations,

$$\mathbb{E}\left[\mathbf{x}_t \mathbf{w}'_{t+1}\right] = \mathbb{E}\left[\mathbb{E}_t[\mathbf{x}_t \mathbf{w}'_{t+1}]\right] = \mathbb{E}\left[\mathbf{x}_t \mathbb{E}_t[\mathbf{w}'_{t+1}]\right]$$

Since $\{\mathbf{w}_t\}$ is an MDS, we have $\mathbb{E}_t[\mathbf{w}'_{t+1}] = \mathbb{E}_t[\mathbf{w}_{t+1}]' = \mathbf{0}'$

It follows that $\mathbb{E}\left[\mathbf{x}_t \mathbf{w}'_{t+1}\right] = \mathbb{E}\left[\mathbf{0}\right] = \mathbf{0}$

Returning to the dynamics of $\mathbf{\Sigma}_t := \mathrm{var}[\mathbf{x}_t]$, we have

$$\begin{aligned}\mathrm{var}[\mathbf{x}_{t+1}] &= \mathbb{E}\left[(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1})(\mathbf{x}_{t+1} - \boldsymbol{\mu}_{t+1})'\right] \\ &= \mathbb{E}\left[(\mathbf{A}(\mathbf{x}_t - \boldsymbol{\mu}_t) + \mathbf{C}\mathbf{w}_{t+1})(\mathbf{A}(\mathbf{x}_t - \boldsymbol{\mu}_t) + \mathbf{C}\mathbf{w}_{t+1})'\right]\end{aligned}$$

The right hand side is equal (**Ex.**) to

$$\begin{aligned}&\mathbb{E}\left[\mathbf{A}(\mathbf{x}_t - \boldsymbol{\mu}_t)(\mathbf{x}_t - \boldsymbol{\mu}_t)'\mathbf{A}'\right] + \mathbb{E}\left[\mathbf{A}(\mathbf{x}_t - \boldsymbol{\mu}_t)\mathbf{w}_{t+1}'\mathbf{C}'\right] \\ &\qquad + \mathbb{E}\left[\mathbf{C}\mathbf{w}_{t+1}(\mathbf{x}_t - \boldsymbol{\mu}_t)'\mathbf{A}'\right] + \mathbb{E}\left[\mathbf{C}\mathbf{w}_{t+1}\mathbf{w}_{t+1}'\mathbf{C}'\right]\end{aligned}$$

Some further manipulations (**Ex.**) lead to

$$\mathbf{\Sigma}_{t+1} = \mathbf{A}\mathbf{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$$

Matrix Dynamics

Incidentally, the law of motion

$$\mathbf{\Sigma}_{t+1} = \mathbf{A}\mathbf{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$$

is an example of a **matrix difference equation**

We can think of it as a dynamical system (S, g) where

- S is the set of $N \times N$ matrices
- $g(\mathbf{\Sigma}) = \mathbf{A}\mathbf{\Sigma}\mathbf{A}' + \mathbf{C}\mathbf{C}'$ maps S to S

Then $\mathbf{\Sigma}_t = g^t(\mathbf{\Sigma}_0)$

Limits of Moments

As we've seen, the dynamics of the mean vector is given by

$$\boldsymbol{\mu}_{t+1} = \mathbf{A}\boldsymbol{\mu}_t + \mathbf{b} \tag{5}$$

If $\rho(\mathbf{A}) < 1$, then this sequence converges

By our earlier results on non-stochastic systems, the unique steady state is

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

Moreover, by those same results,

$$\boldsymbol{\mu}_t \to \boldsymbol{\mu}^* \quad \text{as} \quad t \to \infty \quad \text{regardless of } \boldsymbol{\mu}_0$$

The variance covariance matrices follow

$$\boldsymbol{\Sigma}_{t+1} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$$

A steady state of this system is a $\boldsymbol{\Sigma}$ satisfying

$$\boldsymbol{\Sigma} = \mathbf{A}\boldsymbol{\Sigma}\mathbf{A}' + \mathbf{C}\mathbf{C}' \tag{6}$$

By the results on Lyapunov equations, a unique solution exists whenever $\rho(\mathbf{A}) < 1$

To summarize, if $\rho(\mathbf{A}) < 1$, then

$$\boldsymbol{\mu}_t \to \boldsymbol{\mu}^* \quad \text{and} \quad \boldsymbol{\Sigma}_t \to \boldsymbol{\Sigma}^*$$

where $\boldsymbol{\mu}^* := \sum_{i=0}^{\infty} \mathbf{A}^i \mathbf{b}$ and $\boldsymbol{\Sigma}^*$ is the unique solution to (6)

We can interpret

- $\boldsymbol{\mu}^*$ as the long run mean of the process
- $\boldsymbol{\Sigma}^*$ as the long run variance-covariance matrix

In particular, if $\mathbf{x}_t$ follows our model

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1} \tag{7}$$

then

1. $\mathbb{E}\left[\mathbf{x}_t\right] = \boldsymbol{\mu}^* \implies \mathbb{E}\left[\mathbf{x}_{t+1}\right] = \boldsymbol{\mu}^*$
2. $\mathrm{var}[\mathbf{x}_t] = \boldsymbol{\Sigma}^* \implies \mathrm{var}[\mathbf{x}_{t+1}] = \boldsymbol{\Sigma}^*$

Ex. Check this directly using (7) and the information about $\boldsymbol{\mu}^*$ and $\boldsymbol{\Sigma}^*$ on the previous slide

Example. Let's see this in the scalar case, where

$$x_{t+1} = ax_t + b + cw_{t+1} \quad \text{with} \quad \{w_t\} \overset{\text{IID}}{\sim} N(0,1)$$

Our results tell us that the long run mean is $\boldsymbol{\mu}^* := \sum_{i=0}^{\infty} \mathbf{A}^i \mathbf{b}$

In the scalar case this is just

$$\mu^* := \frac{b}{1-a}$$

So if $|a| < 1$ we should expect that

$$\mu_t := \mathbb{E}[x_t] \to \frac{b}{1-a} \quad \text{as} \quad t \to \infty$$

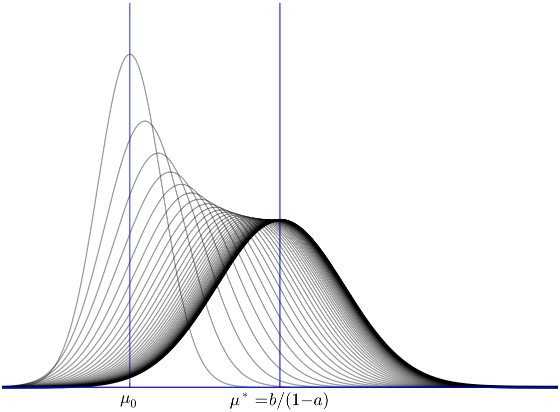

Figure : Convergence of μ_t to μ^* in the scalar model

Dynamics of Marginal Distributions

We've now learned to track $\mathbb{E}[\mathbf{x}_t]$ and $\mathrm{var}[\mathbf{x}_t]$

This gives us some information as to

1. where probability mass is centered
2. how spread out it is, etc.

But it's not as good as knowing all probabilities

That is, it's not as good as knowing the full distribution of $\mathbf{x}_t$

Typically this is a hard problem

... Unless the shocks are normally distributed

So let's consider again the model

$$\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{b} + \mathbf{C}\mathbf{w}_{t+1}$$

Previously we assumed that $\{\mathbf{w}_t\}$ is an MDS

Now we strengthen this to

$$\{\mathbf{w}_t\} \overset{\text{IID}}{\sim} N(\mathbf{0}, \mathbf{I})$$

Fact. Under these assumptions,

1. $\mathbf{x}_0$ constant $\implies$ $\mathbf{x}_t$ is normally distributed for all t
2. $\mathbf{x}_0$ normally distributed $\implies$ $\mathbf{x}_t$ is normally distributed for all t

Proof of Normality

Our model has MA representation

$$\mathbf{x}_t = \mathbf{A}^t \mathbf{x}_0 + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{b} + \sum_{i=0}^{t-1} \mathbf{A}^i \mathbf{C} \mathbf{w}_{t-i}$$

Since

1. $\mathbf{w}_t$ is normally distributed for all t
2. linear operations on normal RVs produce normal RVs,
3. $\mathbf{x}_0$ is constant or normal

it follows that $\mathbf{x}_t$ is normal

The Distribution of $\mathbf{x}_t$ Under Normality

Recall that $\{\mathbf{w}_t\} \stackrel{\text{IID}}{\sim} N(\mathbf{0}, \mathbf{I})$ is a special case of an MDS

Hence our earlier results on moments are still valid:

$$\boldsymbol{\mu}_{t+1} = \mathbf{A}\boldsymbol{\mu}_t + \mathbf{b} \quad \text{and} \quad \boldsymbol{\Sigma}_{t+1} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$$

Initial conditions are the mean and variance of $\mathbf{x}_0$

Since $\mathbf{x}_t$ is normal it follows that

$$\mathbf{x}_t \sim N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t) \quad \text{for all } t$$

This is a complete description of distribution dynamics for $\{\mathbf{x}_t\}$

Example. Let $\mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{w}_{t+1}$ with $\{\mathbf{w}_t\} \overset{\text{IID}}{\sim} N(\mathbf{0}, \mathbf{I})$

Suppose that $\mathbf{x}_0$ is a constant

Using our rule $\boldsymbol{\mu}_{t+1} = \mathbf{A}\boldsymbol{\mu}_t + \mathbf{b}$ for calculating the mean we have

$$\boldsymbol{\mu}_t = \boldsymbol{\mu}_{t-1} = \cdots = \boldsymbol{\mu}_0 = \mathbf{x}_0$$

The dynamics $\boldsymbol{\Sigma}_{t+1} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$ becomes

$$\boldsymbol{\Sigma}_{t+1} = \boldsymbol{\Sigma}_t + \mathbf{I} \quad \text{with} \quad \boldsymbol{\Sigma}_0 = \mathbf{0}$$

Thus,

$$\mathbf{x}_t \sim N(\mathbf{x}_0, t\mathbf{I}) \quad \text{where} \quad t\mathbf{I} = \operatorname{diag}(t, t, \ldots, t)$$

This process is called a Gaussian random walk

Example. Let $\mathbf{x}_{t+1} = \mathbf{A}\mathbf{x}_t + \mathbf{C}\mathbf{w}_{t+1}$ with

1. $\{\mathbf{w}_t\} \overset{\text{IID}}{\sim} N(\mathbf{0}, \mathbf{I})$
2. $\mathbf{x}_0$ constant and

$$\mathbf{x}_0 = \begin{pmatrix} 1.5 \\ -1.1 \end{pmatrix}$$

3. $\mathbf{A}$ and $\mathbf{C}$ have values

$$\mathbf{A} = \begin{pmatrix} 0.6 & -0.7 \\ 0.6 & 0.65 \end{pmatrix} \qquad \mathbf{C} = \begin{pmatrix} 0.1 & 0 \\ 0 & 0.1 \end{pmatrix}$$

We can use the rules

$$\boldsymbol{\mu}_{t+1} = \mathbf{A}\boldsymbol{\mu}_t + \mathbf{b} \quad \text{and} \quad \boldsymbol{\Sigma}_{t+1} = \mathbf{A}\boldsymbol{\Sigma}_t\mathbf{A}' + \mathbf{C}\mathbf{C}'$$

to track the distribution dynamics on a computer

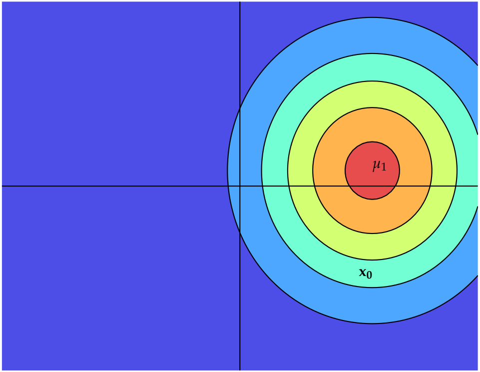

Figure : The density $N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ at $t = 1$

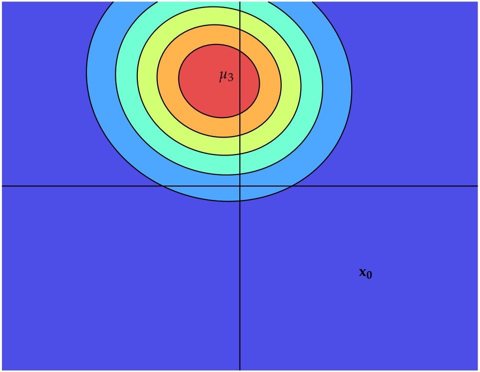

Figure : The density $N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ at $t = 3$

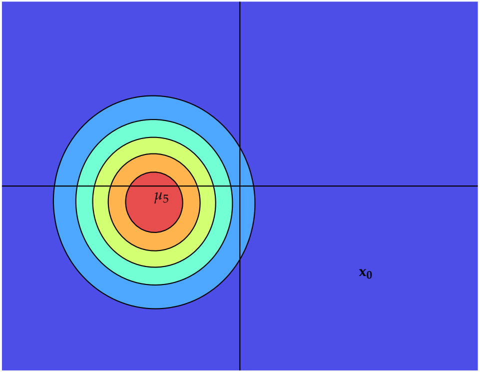

Figure : The density $N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ at $t = 5$

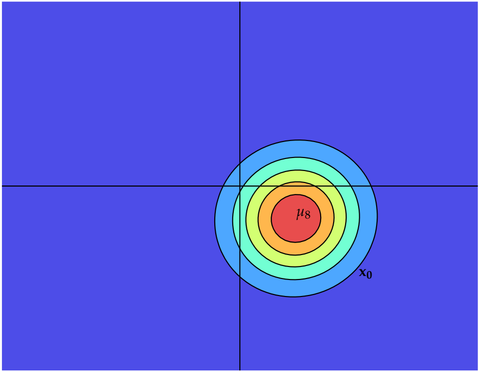

Figure : The density $N(\boldsymbol{\mu}_t, \boldsymbol{\Sigma}_t)$ at $t = 8$