VectorBiTE Methods Training Bayesian State Space Modeling for Time Series Data

The VectorBiTE Team
(John W. Smith, Virginia Tech)

Summer 2021

Prerequisites

Prerequisites

- Familiarity with the R programming language

Prerequisites

- Familiarity with the R programming language
- Rough understanding of Bayesian statistics and Likelihoods

Prerequisites

- Familiarity with the `R` programming language
- Rough understanding of Bayesian statistics and Likelihoods
- Exposure to `JAGS` and related MCMC packages

Why State Space Models?

Why State Space Models?

- The models that we use for physical processes are not perfect

Rogue waves, long thought to be sailor folklore, have been verified and have caused scientists to reconsider their theory behind waves

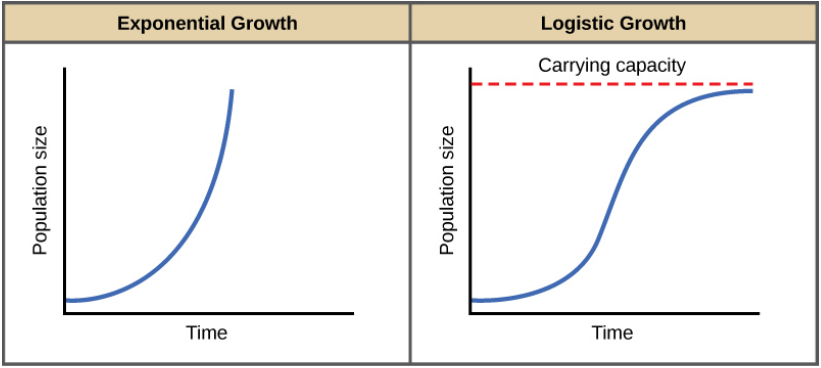
Why State Space Models?

- Ability to make inference on unobserved processes

Temperature is a common variable used as a proxy for Mosquito life traits

Why State Space Models?

- Physical processes are often non-linear and/or non-Gaussian



Population growth is an example of a physical process that is typically non-linear and/or non-Gaussian

What is a State Space Model?

What is a State Space Model?

- Broadly speaking, a state space model is when we have a process x whose value at the current time point only depends on where it was at the last timepoint
  - $x_t \sim x_{t-1}$

What is a State Space Model?

- Broadly speaking, a state space model is when we have a process x whose value at the current time point only depends on where it was at the last timepoint
  - $x_t \sim x_{t-1}$
- x is not observed, instead we observe y which is related to x
  - $y_t \sim x_t$

What is a State Space Model?

- Broadly speaking, a state space model is when we have a process x whose value at the current time point only depends on where it was at the last timepoint
  - $x_t \sim x_{t-1}$
- x is not observed, instead we observe y which is related to x
  - $y_t \sim x_t$
- $y_{1:T}$ are independent of one another conditional on $x_{1:T}$

What is a State Space Model?

What is a State Space Model?

In general, state space models will have the form:

$$x_t = f(x_{t-1}|\Theta)$$
$$y_t = g(x_t|\Theta)$$

What is a State Space Model?

In general, state space models will have the form:

$$x_t = f(x_{t-1}|\Theta)$$
$$y_t = g(x_t|\Theta)$$

- $f(x_t|x_{t-1}, \Theta)$ is called the *evolution function*

What is a State Space Model?

In general, state space models will have the form:

$$x_t = f(x_{t-1}|\Theta)$$
$$y_t = g(x_t|\Theta)$$

- $f(x_t|x_{t-1}, \Theta)$ is called the *evolution function*
- $g(y_t|x_t, \Theta)$ is called the *observation density function*

What is a State Space Model?

In general, state space models will have the form:

$$x_t = f(x_{t-1}|\Theta)$$
$$y_t = g(x_t|\Theta)$$

- $f(x_t|x_{t-1}, \Theta)$ is called the *evolution function*
- $g(y_t|x_t, \Theta)$ is called the *observation density function*
- Θ are the additional parameters that are used in $f(\cdot)$ and $g(\cdot)$

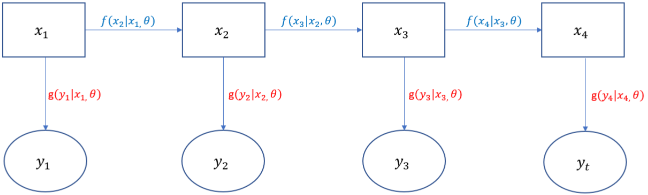
What is a State Space Model?

In general, state space models will have the form:

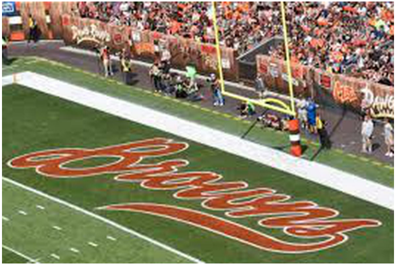
$$x_t = f(x_{t-1}|\Theta)$$
$$y_t = g(x_t|\Theta)$$

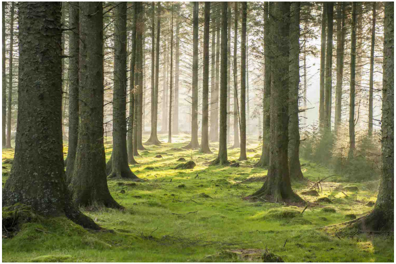
- $f(x_t|x_{t-1}, \Theta)$ is called the *evolution function*
- $g(y_t|x_t, \Theta)$ is called the *observation density function*
- Θ are the additional parameters that are used in $f(\cdot)$ and $g(\cdot)$
- $f(\cdot)$ and $g(\cdot)$ are both stochastic

Illustration



What Can We Use SSMs For?





State Space Models: Example 1

State Space Models: Example 1

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = x_{t-1} + \epsilon_{proc}$$
$$y_t = x_t + \epsilon_{obs}$$
$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

State Space Models: Example 1

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = x_{t-1} + \epsilon_{proc}$$
$$y_t = x_t + \epsilon_{obs}$$
$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$

State Space Models: Example 1

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = x_{t-1} + \epsilon_{proc}$$

$$y_t = x_t + \epsilon_{obs}$$

$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = x_{t-1}, \phi^* = \phi)$

State Space Models: Example 1

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = x_{t-1} + \epsilon_{proc}$$

$$y_t = x_t + \epsilon_{obs}$$

$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = x_{t-1}, \phi^* = \phi)$
- $g(y_t|x_t, \Theta) \sim N(\mu^* = x_t, \phi^* = \tau)$

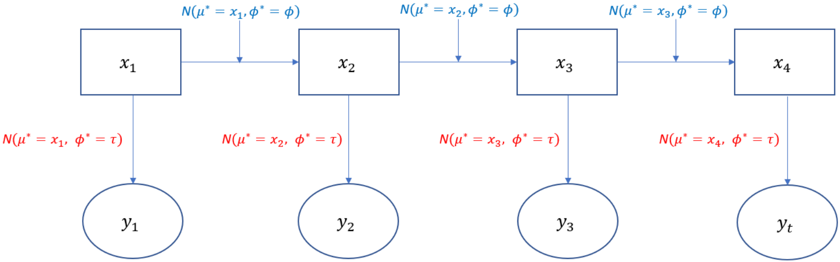
State Space Models: Example 1

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

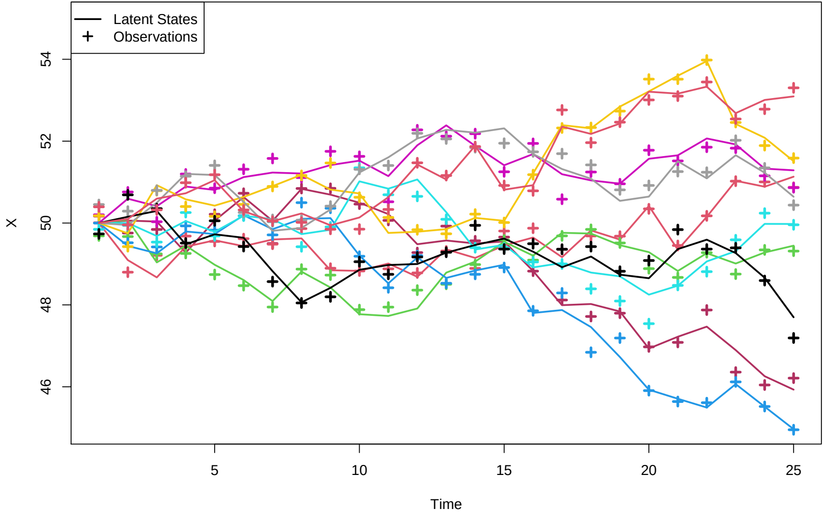
$$x_t = x_{t-1} + \epsilon_{proc}$$
$$y_t = x_t + \epsilon_{obs}$$
$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = x_{t-1}, \phi^* = \phi)$
- $g(y_t|x_t, \Theta) \sim N(\mu^* = x_t, \phi^* = \tau)$
- This is a special case of the State Space Model called a *Linear Gaussian State Space model*, or *Normal Dynamic Linear model* (NDLM)

State Space Models: Example 1



State Space Models: Example 1



State Space Models: Example 2

State Space Models: Example 2

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$
\begin{aligned}
x_t &= \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t) + \epsilon_{proc} \\
y_t &= \frac{x_t^2}{20} + \epsilon_{obs} \\
&\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)
\end{aligned}
$$

State Space Models: Example 2

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1},\Theta)$, and $g(y_t|x_t,\Theta)$.

$$x_t = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t) + \epsilon_{proc}$$

$$y_t = \frac{x_t^2}{20} + \epsilon_{obs}$$

$$\epsilon_{proc} \sim N(0,\phi), \epsilon_{obs} \sim N(0,\tau)$$

- $\Theta = \{\phi, \tau\}$

State Space Models: Example 2

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t) + \epsilon_{proc}$$

$$y_t = \frac{x_t^2}{20} + \epsilon_{obs}$$

$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t), \phi^* = \phi)$

State Space Models: Example 2

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1 + x_{t-1}^2} + 8\cos(1.2t) + \epsilon_{proc}$$

$$y_t = \frac{x_t^2}{20} + \epsilon_{obs}$$

$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t), \phi^* = \phi)$
- $g(y_t|x_t, \Theta) \sim N(\mu^* = \frac{x_t^2}{20}, \phi^* = \tau)$

State Space Models: Example 2

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

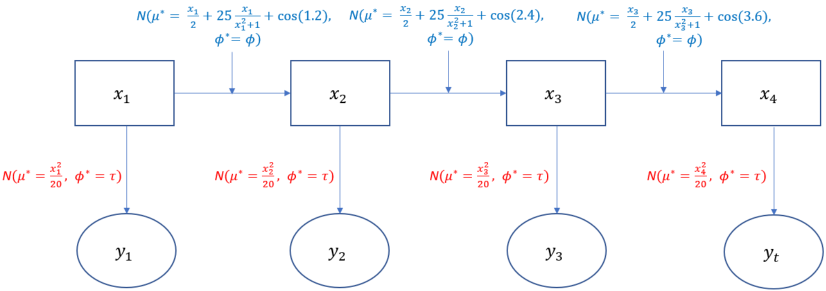
$$x_t = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1 + x_{t-1}^2} + 8\cos(1.2t) + \epsilon_{proc}$$

$$y_t = \frac{x_t^2}{20} + \epsilon_{obs}$$

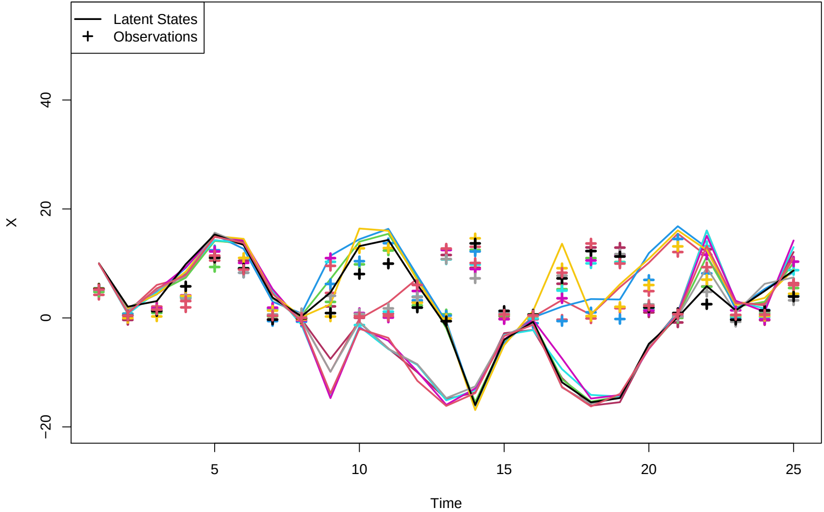
$$\epsilon_{proc} \sim N(0, \phi), \epsilon_{obs} \sim N(0, \tau)$$

- $\Theta = \{\phi, \tau\}$
- $f(x_t|x_{t-1}, \Theta) \sim N(\mu^* = \frac{x_{t-1}}{2} + 25\frac{x_{t-1}}{1+x_{t-1}^2} + 8\cos(1.2t), \phi^* = \phi)$
- $g(y_t|x_t, \Theta) \sim N(\mu^* = \frac{x_t^2}{20}, \phi^* = \tau)$
- This is an example of a *Gaussian State Space model*. It differs from Example 1 because it is not linear (why?)

State Space Models: Example 2



State Space Models: Example 2



State Space Models: Example 3

State Space Models: Example 3

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$$
$$y_t \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$$

State Space Models: Example 3

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$$

$$y_t \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$$

- $\Theta = \{a, \sigma^2, b, \tau^2\}$

State Space Models: Example 3

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$$
$$y_t \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$$

- $\Theta = \{a, \sigma^2, b, \tau^2\}$
- $f(x_t|x_{t-1}, \Theta) \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$

State Space Models: Example 3

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

$$x_t \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$$

$$y_t \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$$

- $\Theta = \{a, \sigma^2, b, \tau^2\}$
- $f(x_t|x_{t-1}, \Theta) \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$
- $g(y_t|x_t, \Theta) \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$

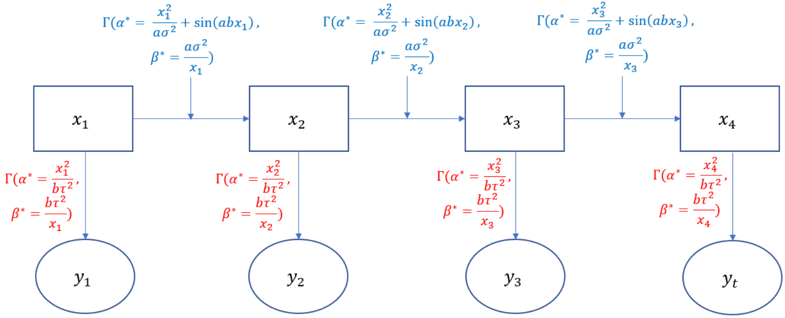
State Space Models: Example 3

Let's consider the following example, and identify the corresponding $\Theta, f(x_t|x_{t-1}, \Theta)$, and $g(y_t|x_t, \Theta)$.

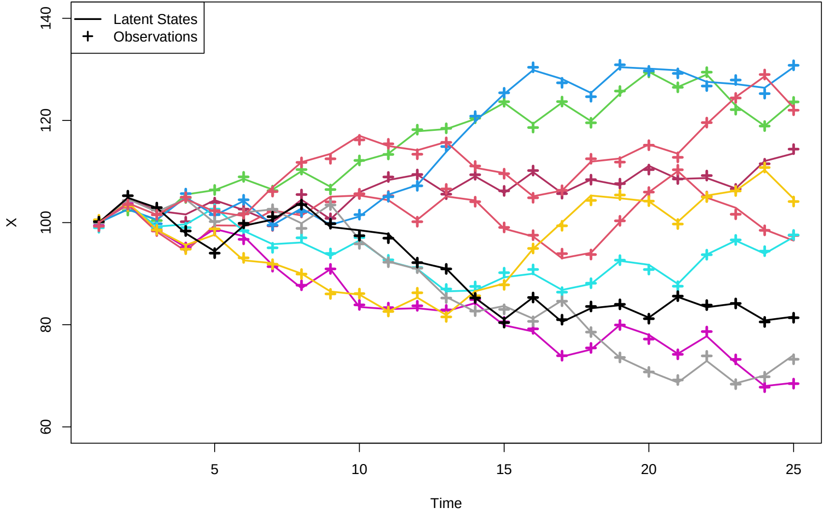
$$x_t \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$$
$$y_t \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$$

- $\Theta = \{a, \sigma^2, b, \tau^2\}$
- $f(x_t|x_{t-1}, \Theta) \sim \Gamma(\alpha^* = \frac{x_{t-1}^2}{a\sigma^2} + \sin(abx), \beta^* = \frac{a\sigma^2}{x_{t-1}})$
- $g(y_t|x_t, \Theta) \sim \Gamma(\alpha^* = \frac{x_t^2}{b\tau^2}, \beta^* = \frac{b\tau^2}{x_t})$
- This is an example of a *Non-linear non-Gaussian State Space model*

State Space Models: Example 3



State Space Models: Example 3



So What's the Catch?

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated
  - Imagine a simple time series regression with 10 data points and 1 parameter to be estimated

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated
  - Imagine a simple time series regression with 10 data points and 1 parameter to be estimated
  - If we fit this as a state space model, suddenly we are estimating 11 parameters

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated
  - Imagine a simple time series regression with 10 data points and 1 parameter to be estimated
  - If we fit this as a state space model, suddenly we are estimating 11 parameters
  - This leads to increased computation times

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated
  - Imagine a simple time series regression with 10 data points and 1 parameter to be estimated
  - If we fit this as a state space model, suddenly we are estimating 11 parameters
  - This leads to increased computation times
- $g(x_t|x_{t-1}, \Theta)$ and $f(y_t|x_t, \Theta)$ need to be specified

So What's the Catch?

- In the 3 examples where we looked at potential structures for state space models, we've seen that they can handle basically anything
- Powerful, flexible methods like these often come with drawbacks, so what are they?
- Latent states are parameters that themselves need to be estimated
  - Imagine a simple time series regression with 10 data points and 1 parameter to be estimated
  - If we fit this as a state space model, suddenly we are estimating 11 parameters
  - This leads to increased computation times
- $g(x_t|x_{t-1}, \Theta)$ and $f(y_t|x_t, \Theta)$ need to be specified
  - Though not all parameters in Θ need to be known, we need to know the form of the distribution for the evolution and observation density functions

End of Presentation 1