# Chapter 3 - Temporal processes

Odd Kolbjørnsen and Geir Storvik

January 23 2017

# Temporal processes

Data collected over time

- Past, present, future, change
- Temporal aspect important?

Two (separate) approaches to modelling

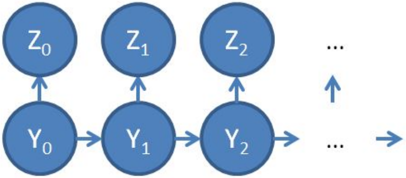
- Statistical
  - Variability through randomness
  - Learning dynamic structure from data
- Dynamic system theory
  - Typically deterministic

# Hierarchical (statistical) models

- Data model
- Process model

In time series setting - called *state space models*
Example

$$
\begin{aligned}
Y_1 \sim & N(\mu_1, \sigma_1^2) \\
Y_t = & \alpha Y_{t-1} + W_t, t = 2, 3, ... & \qquad W_t \stackrel{ind}{\sim} (0, \sigma_W^2) \\
Z_t = & \beta Y_t + \eta_t, t = 1, 2, 3, ... & \qquad \eta_t \stackrel{ind}{\sim} (0, \sigma_\eta^2)
\end{aligned}
$$



# Characterization of temporal processes

Denoted $\mathbf{Y}(\cdot)$ or $\{\mathbf{Y}(r) : r \in D_t\}$

- $r$ time index
- $\mathbf{Y}(r)$ possibly multivariate, deterministic or stochastic
- $D_t \in \mathcal{R}^1$

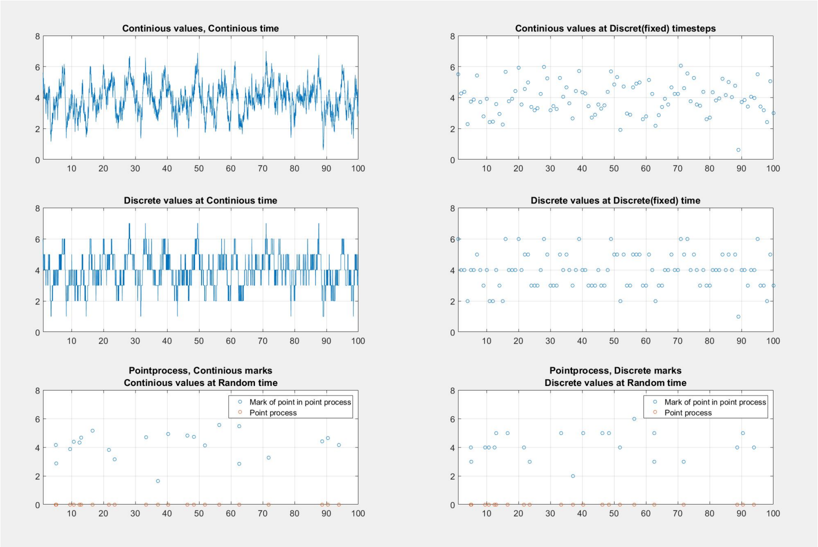
$D_t$ specify type of process

- Continuous time process: $D_t \in (-\infty, \infty)$ or $[0, \infty)$: $\mathbf{Y}(t)$
- Discrete time process: $D_t \in \mathcal{N}$ or $\mathcal{N}^+$: $\mathbf{Y}_t$
- Point process: Random times
  Example: Time of tornado. Mark: Severity of tornado

Continuous time:

- Measured discrete
- Modelled discrete?
- Dynamic system theory: Often modelled continuous (PDE's)

# Temporal processes, Discrete values and time



# Joint and conditional distributions

Time series: $\{Y_t : t = 0, ..., T\}$.
Joint distribution

$$[Y_0, Y_1, ..., Y_T] = [Y_T | Y_{T-1}, ..., Y_0][Y_{T-1} | Y_{T-2}, ..., Y_0] \ldots [Y_1 | Y_0][Y_0]$$

*Markov assumption*

$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}]$$

imply

$$[Y_0, Y_1, ..., Y_T] = [Y_0] \prod_{t=1}^{T} [Y_t | Y_{t-1}]$$

Dramatic simplification!
Possible extension:

$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}, \ldots, Y_{t-k}]$$

# Joint and conditional distributions

Time series: $\{Y_t : t = 0, ..., T\}$.
Joint distribution

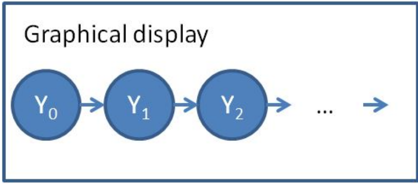
$$[Y_0, Y_1, ..., Y_T] = [Y_T | Y_{T-1}, ..., Y_0][Y_{T-1} | Y_{T-2}, ..., Y_0] \dots [Y_1 | Y_0][Y_0]$$

*Markov assumption*

$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}]$$

imply

$$[Y_0, Y_1, ..., Y_T] = [Y_0] \prod_{t=1}^{T} [Y_t | Y_{t-1}]$$




Dramatic simplification!
Possible extension:

$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}, \dots, Y_{t-k}]$$

# Joint and conditional distributions

Time series: $\{Y_t : t = 0, ..., T\}$.
Joint distribution

$$[Y_0, Y_1, ..., Y_T] = [Y_T | Y_{T-1}, ..., Y_0][Y_{T-1} | Y_{T-2}, ..., Y_0] \ldots [Y_1 | Y_0][Y_0]$$

*Markov assumption*

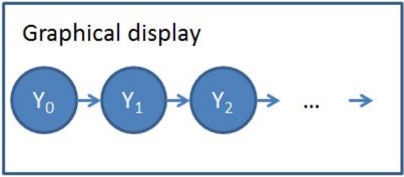
$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}]$$

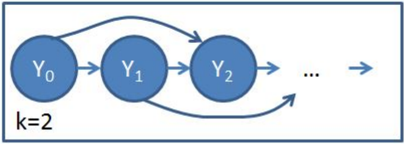
imply

$$[Y_0, Y_1, ..., Y_T] = [Y_0] \prod_{t=1}^{T} [Y_t | Y_{t-1}]$$

Dramatic simplification!
Possible extension:

$$[Y_t | Y_{t-1}, ..., Y_0] = [Y_t | Y_{t-1}, \ldots, Y_{t-k}]$$



# Deterministic and stochastic modelling

PDE:

$$\frac{dY(t)}{dt} = f(Y(t)), \quad t \geq 0$$

Completely determined by $Y(0)$.
Discrete time:

$$Y_t = \mathcal{M}(Y_{t-1}), \quad t = 1, 2, ...$$

Completely determined by $Y_0$.
Extension to stochastic models:

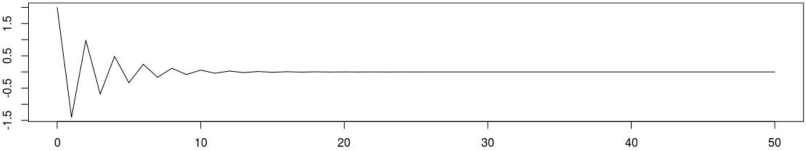
$$Y_t = \mathcal{M}(Y_{t-1}) + \eta_t, \quad t = 1, 2, ...$$

No clear distinction

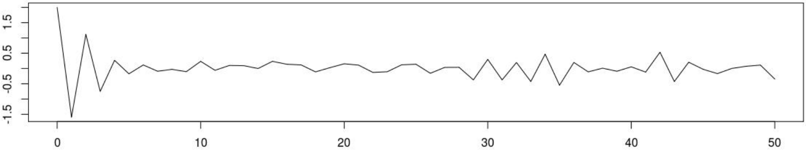
- Deterministic *chaotic* processes can look random

# Deterministic dynamical systems

Simple difference model of order 1: $\alpha_t = \theta_1 \alpha_{t-1}$.
Homogeneous system
Example: $\alpha_0 = 1, \theta_1 = -0.7$. Equilibrium $= 0$




Nonhomogeneous system: $\alpha_t = \theta_1 \alpha_{t-1} + g_t$.
Possible equilibrium point depend on $\{g_t\}$.
Forces external to system can be modelled as a random process.
Example: $g_t \sim N(0, 0.2^2)$.



# Backwards-shift operator

$B$: Backwards-shift operator: $B\alpha_t = \alpha_{t-1}$

$$\alpha_t = \theta_1\alpha_{t-1} = \theta_1 B\alpha_t$$

Extensions:

$$\alpha_t = \theta_1\alpha_{t-1} + \theta_2\alpha_{t-2} = \theta_1 B\alpha_t + \theta_2 B^2\alpha_t$$

Give

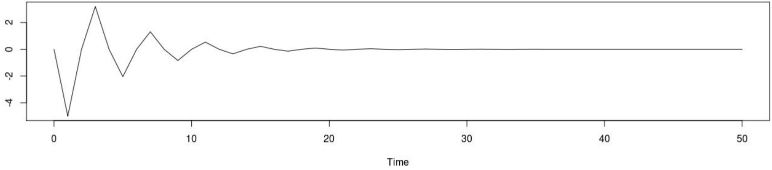
$$(1 - \theta_1 B - \theta_2 B^2)\alpha_t = 0$$

$(1 - \theta_1 z - \theta_2 z^2)$ is the *characteristic polynomial*
Behavior of dynamical system is described by the roots of this polynomial

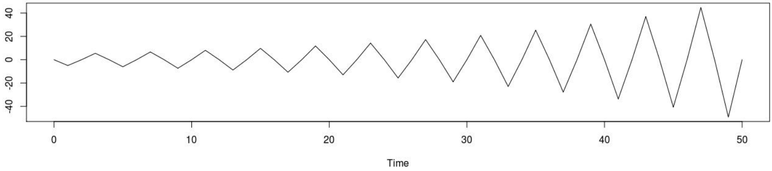
- Roots $|z_1| > 1$ and $|z_2| > 1$: Equilibrium value is attracting (stable)
- Roots $|z_1| < 1$ or $|z_2| < 1$: Equilibrium value is repelling (unstable)

# Examples

Dynamical system $\alpha_t = -0.64\alpha_{t-2}, \quad t = 2, 3, ..., \alpha_0 = 0, \alpha_1 = -5$
Roots: $z_1 = 1.25i, z_2 = -1.25i$




Dynamical system $\alpha_t = -1.1\alpha_{t-2}, \quad t = 2, 3, ..., \alpha_0 = 0, \alpha_1 = -5$
Roots: $z_1 = 0.953, z_2 = -0.953$



# Multivariate time series

$\boldsymbol{\alpha}_t = (\alpha_t(1), ..., \alpha_t(p_\alpha))^T$
$1, ..., p_\alpha$ might be spatial locations!
Generalization of AR(1) model

$$\boldsymbol{\alpha}_t = \mathbf{M}\boldsymbol{\alpha}_{t-1} = \mathbf{M}^t\boldsymbol{\alpha}_0$$

Note: AR(3)=Multivariate AR(1):

$$\begin{aligned}
\beta_t =& \theta_1\beta_{t-1} + \theta_2\beta_{t-2} + \theta_3\beta_{t-3} \\
\begin{pmatrix}\beta_t \\ \beta_{t-1} \\ \beta_{t-2}\end{pmatrix} =& \begin{pmatrix}\theta_1 & \theta_2 & \theta_3 \\ 1 & 0 & 0 \\ 0 & 1 & 0\end{pmatrix} \begin{pmatrix}\beta_{t-1} \\ \beta_{t-2} \\ \beta_{t-3}\end{pmatrix} \\
\boldsymbol{\alpha}_t =& \mathbf{M}\boldsymbol{\alpha}_{t-1}
\end{aligned}$$

# Spectral representation of linear dynamic models

Assume $\boldsymbol{\alpha}_t = \mathbf{M}\boldsymbol{\alpha}_{t-1} = \mathbf{M}^t\boldsymbol{\alpha}_0$
Linear algebra:

- $\mathbf{M} = \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{-1}$
- $\boldsymbol{\Lambda}$: Diagonal matrix with eigenvalues
- $\mathbf{W}$: Right eigenvectors

$$
\begin{aligned}
\mathbf{M}^2 =& \mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{-1}\mathbf{W}\boldsymbol{\Lambda}\mathbf{W}^{-1} \\
=& \mathbf{W}\boldsymbol{\Lambda}^2\mathbf{W}^{-1} \\
\mathbf{M}^t =& \mathbf{W}\boldsymbol{\Lambda}^t\mathbf{W}^{-1} \\
\boldsymbol{\alpha}_t =& \mathbf{W}\boldsymbol{\Lambda}^t\mathbf{W}^{-1}\boldsymbol{\alpha}_0 \\
=& \mathbf{W}\boldsymbol{\Lambda}^t\mathbf{c} = \sum_{i=1}^{p_\alpha} c_i\lambda_i^t\mathbf{w}_i
\end{aligned}
$$

Dynamic behaviour depend on $\lambda_i$'s

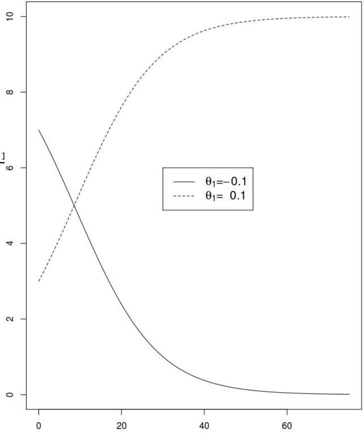
# Nonlinear discrete dynamical systems

Example: *Logistic dynamical system*

$$
\begin{aligned}
\alpha_t =& \mathcal{M}(\alpha_{t-1}) \\
=& (1+\theta_1)\alpha_{t-1} - \tfrac{\theta_1}{\theta_0}\alpha_{t-1}^2 \\
=& \alpha_{t-1} + \theta_1\alpha_{t-1}\left[1 - \tfrac{\alpha_{t-1}}{\theta_0}\right] \\
\alpha_t - \alpha_{t-1} =& \theta_1\alpha_{t-1}\left[1 - \tfrac{\alpha_{t-1}}{\theta_0}\right]
\end{aligned}
$$

Population dynamics:

- $\theta_0$ is carying capacity
- $\theta_1$ is a growth parameter




Small changes in parameters can lead to large changes in equilibrium

# Other issues

- Bifurcation: Small changes in parameters give different sets of equilibrium points
- Chaos behaviour: Sensitivity to initial behaviour
- State space reconstruction: If $\boldsymbol{\alpha}_t = \mathcal{M}(\boldsymbol{\alpha}_{t-1}; \boldsymbol{\theta})$, sometimes possible to reconstruct $\boldsymbol{\alpha}_t$ from lowdimensional observations of $\boldsymbol{\alpha}_t$

# Need for statistical models

- Uncertainty breeds “chaos”
- Sometimes possible to model “chaos” through deterministic systems, but statistical model usually easier
- Statistical models specify structure in behaviour through probability models
- Statistical models can predict, chaotic deterministic systems can not

# Time series

Assume $\{Y_t, t = 0, 1, ...\}$, time series (statistical dynamical model)

Mean function $\mu_t = E(Y_t)$, deterministic

Autocovariance function

$$C_Y(t, r) = \text{cov}(Y_t, Y_r) = E\{(Y_t - \mu_t)(Y_r - \mu_r)\}, \quad t, r \in D_t$$

Variance function $\sigma_t^2 = \text{var}(Y_t) = C_Y(t, t)$

Autocorrelation function

$$\rho_Y(t, r) = \frac{C_Y(t, r)}{\sqrt{C_Y(t, t) C_Y(r, r)}} \in [-1, 1]$$

# Covariance functions are nonnegative definite functions

Consider $V = \sum_{i=1}^{N} a_i Y_{t_i} = \mathbf{a}^T \mathbf{Y}_\mathbf{I}$ for $\mathbf{I} = \{t_i\}_{i=1}^N$ any finite subset of $D_t$

- $\text{var}(V) \geq 0$

$$\begin{aligned}\text{var}[V] =& \sum_{i=1}^{N} a_i^2 \text{var}(Y_{t_i}) + 2 \sum_{1 \leq i < j \leq N} a_i a_j \text{cov}(Y_{t_i}, Y_{t_j}) \\ =& \mathbf{a}^T \mathbf{C}_{\mathbf{Y}_\mathbf{I}} \mathbf{a}\end{aligned}$$

  where $\mathbf{C}_{\mathbf{Y}_I}$ is covariance matrix for $\mathbf{Y}_\mathbf{I}$
- Need $\mathbf{a}^T \mathbf{C}_{\mathbf{Y}_I} \mathbf{a} \geq 0$ for all $\mathbf{a}$
- $\mathbf{C}_{\mathbf{Y}_I}$ specified through $C_Y(\cdot)$
- A function satisfying this requirement is a *nonnegative definite function*

# Stationarity

Weak/second order stationarity require

1. $E(Y_t) = \mu$, for all $t \in D_t$
2. $\text{cov}(Y_t, Y_r) = C_Y(t - r)$ for all $t, r \in D_t$

Strong stationarity For any $(t_1, ..., t_m)$ and any $\tau$

$$[Y_{t_1}, ..., Y_{t_m}] = [Y_{t_1+\tau}, ..., Y_{t_m+\tau}]$$

# Estimation

Assume second-order stationarity

$$\hat{\mu} = \frac{1}{T}\sum_{t=1}^{T} Y_t \qquad \text{Sample mean}$$

$$\widehat{C}(\tau) = \frac{1}{T}\sum_{t=1}^{T-\tau} (Y_{t+\tau} - \hat{\mu})(Y_t - \hat{\mu}) \qquad \text{Empirical autocovariance function}$$

Can divided by $T - \tau$ for unbiased estimate

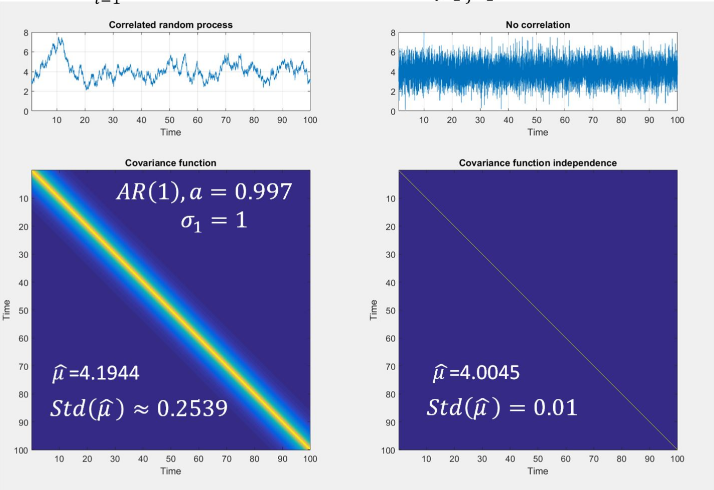
Dividing by $T$ gives a *nonnegative definite function*

$$\hat{\rho}(\tau) = \frac{\widehat{C}_Y(\tau)}{\widehat{C}_Y(0)} \in [-1, 1]$$

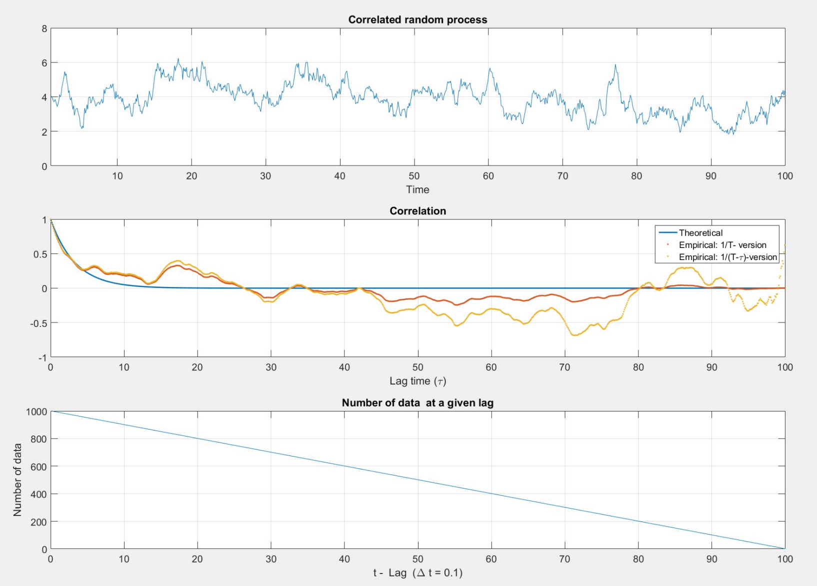
# Example: estimation mean

10 000 samples in intervall

$$\hat{\mu} = \frac{1}{T}\sum_{i=1}^{T} Y_i = \frac{1}{T}\mathbf{1}^T\boldsymbol{Y} \qquad Var(\hat{\mu}) = \frac{1}{T^2}\sum_{i=1}^{T}\sum_{j=1}^{T} Cov(Y_i, Y_j) = \frac{1}{T^2}\mathbf{1}^T\boldsymbol{\Sigma}\mathbf{1}$$



# Example: estimation correlation



# Basic time series

$\{W_t, t \in \mathcal{N}\}$ iid zero-mean random variables with $\text{var}(W_t) = \sigma_w^2$

White noise $Y_t = W_t$

$$C(\tau) = \begin{cases} \sigma_\omega^2, & \tau = 0 \\ 0, & \tau = \pm 1, \pm 2, \cdots \end{cases}$$

Random-walk $Y_t = Y_{t-1} + W_t$. Not stationary

Autoregressive AR(p) $Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + W_t$

Moving average MA(q) $Y_t = W_t + \beta_1 W_{t-1} + \cdots + \beta_q W_{t-q}$

ARMA(p,q) $Y_t = \alpha_1 Y_{t-1} + \cdots + \alpha_p Y_{t-p} + W_t + \beta_1 W_{t-1} + \cdots + \beta_q W_{t-q}$

VAR $\mathbf{Y}_t = \mathbf{M}_1 \mathbf{Y}_{t-1} + \cdots + \mathbf{M}_p \mathbf{Y}_{t-p} + \mathbf{W}_t$

# Parameter estimation

- Method-of-moments
  - Set of linear equations for AR models
- Maximum likelihood
  - Kalman filtering for calculation of likelihood
  - Numerical optimization for maximizing likelihood

# Nonlinear time series

- Typically combination of basic time series models and deterministic dynamical models

$$Y_t = \mathcal{M}(Y_{t-1}, ..., Y_{t-p}) + \tau(Y_{t-1}, ..., Y_{t-p})W_t$$

# Covariance functions for AR/MA processes

- AR(1): $C_Y(\tau) = \alpha_1^\tau$
- MA(1): $C_Y(\tau)$ given by

$$C_Y(\tau) = \sigma_w^2 \begin{cases} 1, & \tau = 0 \\ \beta_1, & \tau = 1 \\ 0, & \tau > 1 \end{cases}$$

General

- AR processes: Decay exponentially
- MA processes: Zero outside lag

# Hierearchichal models

| | Variable | Densities | Notation in book |
|---|---|---|---|
| Data model: | $\mathbf{Z}$ | $p(\mathbf{Z}\|\mathbf{Y},\boldsymbol{\theta})$ | $[\mathbf{Z}\|\mathbf{Y},\boldsymbol{\theta}]$ |
| Process model: | $\mathbf{Y}$ | $p(\mathbf{Y}\|\boldsymbol{\theta})$ | $[\mathbf{Y}\|\boldsymbol{\theta}]$ |
| Parameter model: | $\boldsymbol{\theta}$ | $p(\boldsymbol{\theta})$ | $[\boldsymbol{\theta}]$ |

Time series setting

$$[\mathbf{Z}|\mathbf{Y},\boldsymbol{\theta}]: \quad \mathbf{Z}_t = \mathcal{H}(\mathbf{Y}_t;\boldsymbol{\theta}) + \boldsymbol{\varepsilon}_t \quad \boldsymbol{\varepsilon}_t \sim \text{Gau}(\mathbf{0},\mathbf{R}_t(\boldsymbol{\theta}))$$

$$[\mathbf{Y}|\boldsymbol{\theta}]: \quad \mathbf{Y}_t = \mathcal{M}(\mathbf{Y}_{t-1};\boldsymbol{\theta}) + \boldsymbol{\eta}_t \quad \boldsymbol{\eta}_t \sim \text{Gau}(\mathbf{0},\mathbf{Q}_t(\boldsymbol{\theta}))$$

Of interest:

Filtered estimate: $E(\mathbf{Y}_t|\mathbf{Z}_1,...,\mathbf{Z}_t)$

Forcasted estimate: $E(\mathbf{Y}_{t+\tau}|\mathbf{Z}_1,...,\mathbf{Z}_T)$

Smoothed estimate: $E(\mathbf{Y}_t|\mathbf{Z}_1,...,\mathbf{Z}_T)$