Dynamic Factor Analysis

FISH 507 – Applied Time Series Analysis

Mark Scheuerell

28 January 2021

Topics for today

Deterministic vs stochastic elements

Regression with autocorrelated errors

Regression with temporal random effects

Dynamic Factor Analysis (DFA)

- Forms of covariance matrix
- Constraints for model fitting
- Interpretation of results

A very simple model

Consider this simple model, consisting of a mean μ plus error

$$y_i = \mu + e_i \text{ with } e_i \sim \mathrm{N}(0, \sigma^2)$$

A very simple model

The right-hand side of the equation is composed of *deterministic* and *stochastic* pieces

$$y_i = \underbrace{\mu}_{\text{deterministic}} + \underbrace{e_i}_{\text{stochastic}}$$

A very simple model

Sometime these pieces are referred to as *fixed* and *random*

$$y_i = \underbrace{\mu}_{\text{fixed}} + \underbrace{e_i}_{\text{random}}$$

A very simple model

This can also be seen by rewriting the model

$$y_i = \mu + e_i \text{ with } e_i \sim \mathrm{N}(0, \sigma^2)$$

as

$$y_i \sim \mathrm{N}(\mu, \sigma^2)$$

Simple linear regression

We can expand the deterministic part of the model, as with linear regression

$$y_i = \underbrace{\alpha + \beta x_i}_{\text{mean}} + e_i \text{ with } e_i \sim \mathrm{N}(0, \sigma^2)$$

so

$$y_i \sim \mathrm{N}(\alpha + \beta x_i, \sigma^2)$$

A simple time series model

Consider a simple model with a mean μ plus white noise

$$y_t = \mu + e_t \text{ with } e_t \sim \text{N}(0, \sigma^2)$$

Time series model with covariates

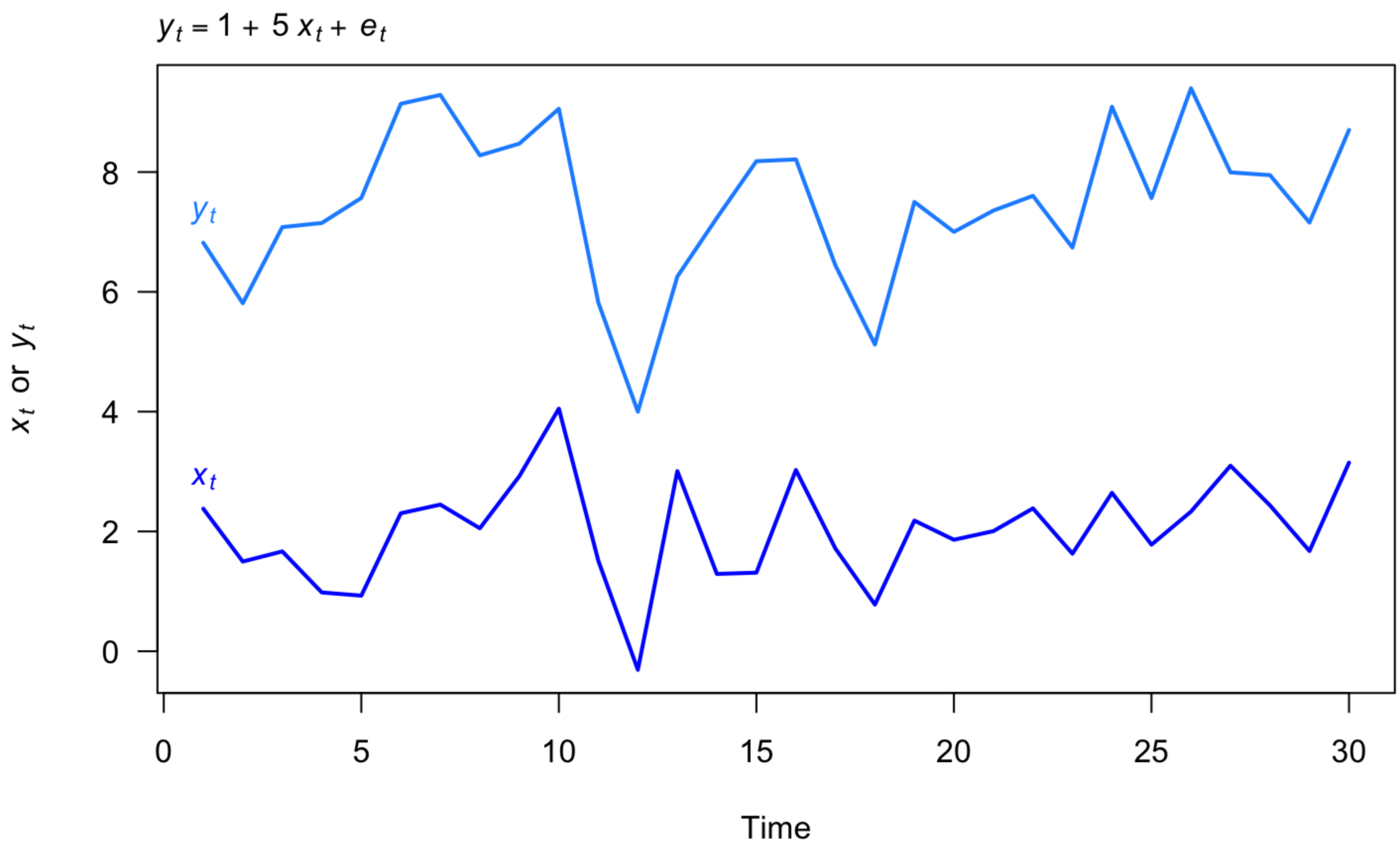
We can expand the deterministic part of the model, as before with linear regression

$$y_t = \underbrace{\alpha + \beta x_t}_{\text{mean}} + e_t \text{ with } e_t \sim \text{N}(0, \sigma^2)$$

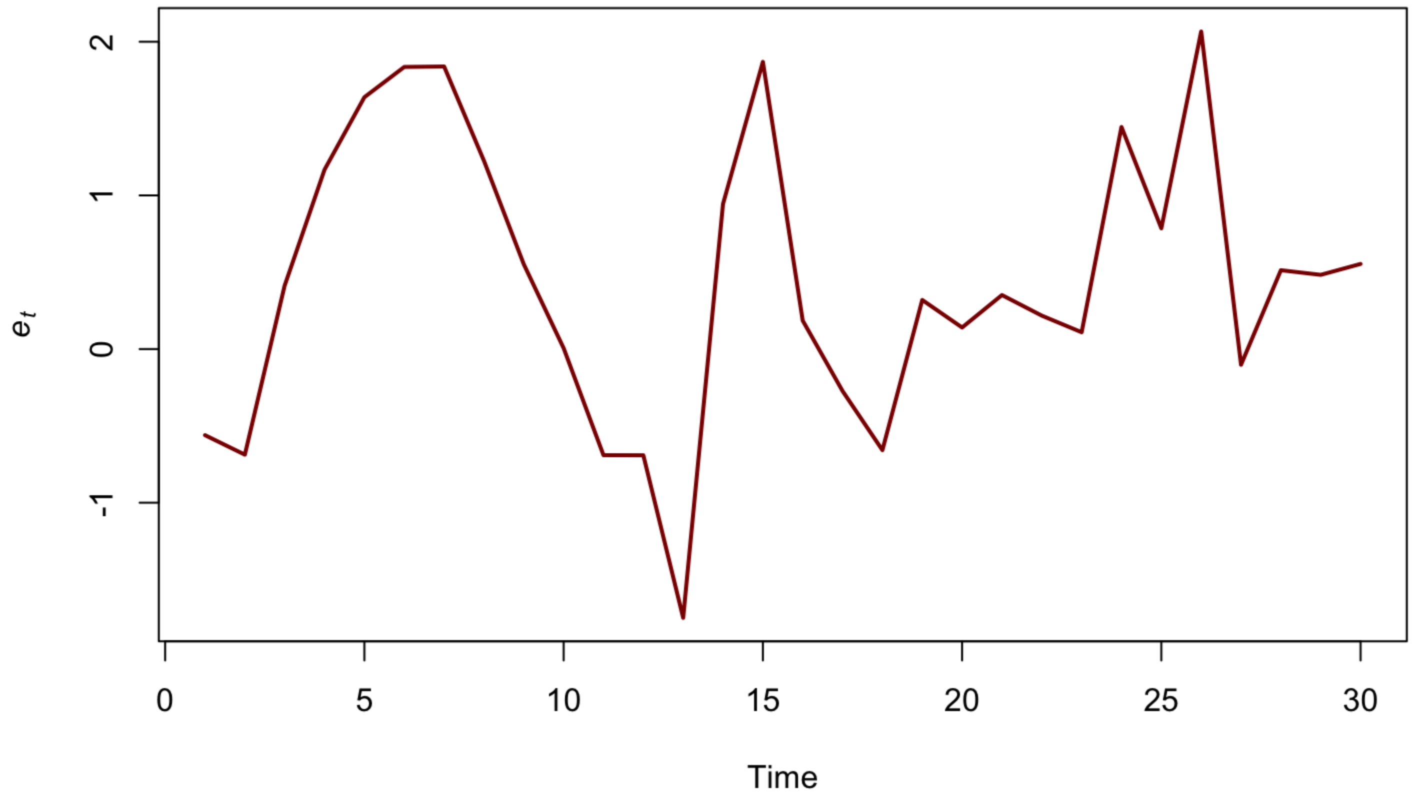
so

$$y_t \sim \text{N}(\alpha + \beta x_t, \sigma^2)$$

Example of linear model

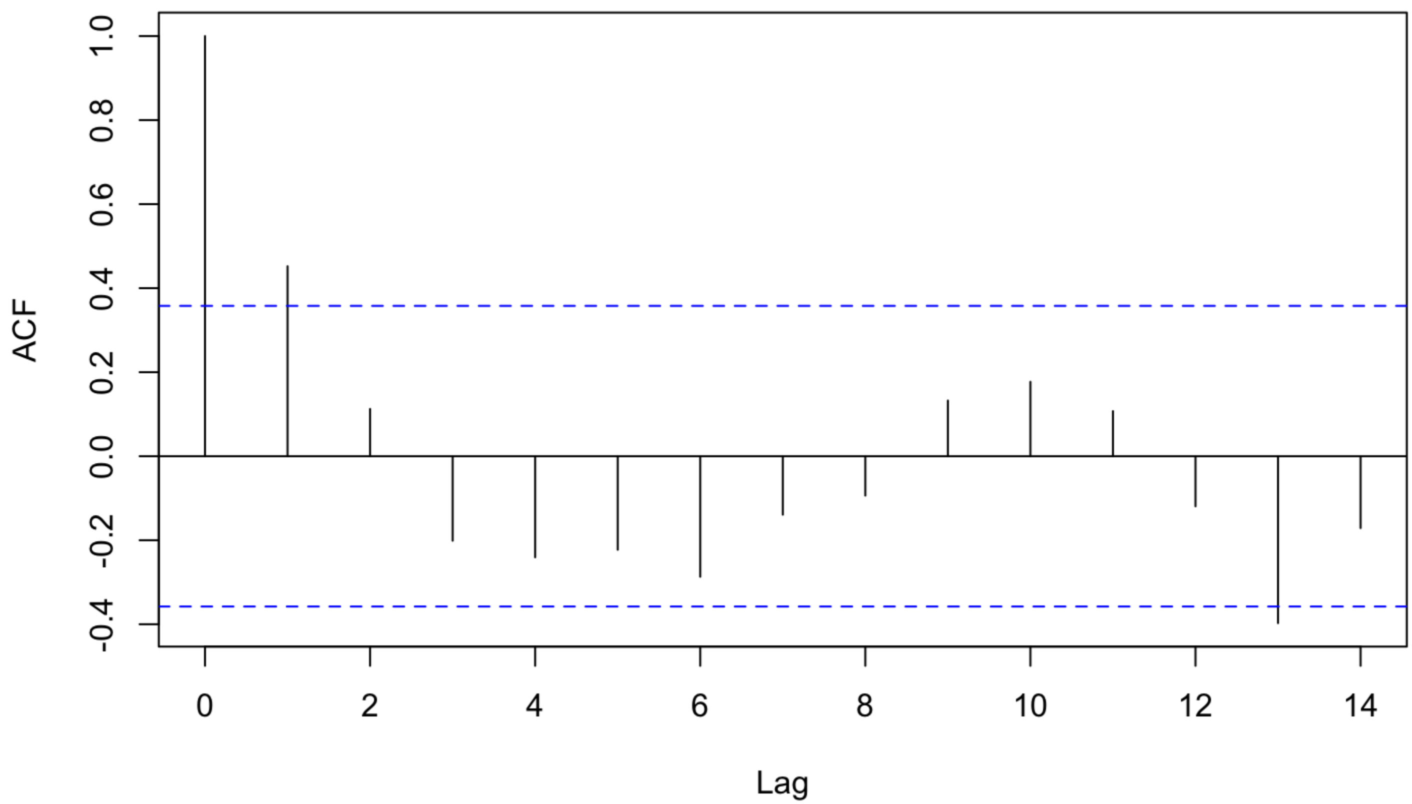


Model residuals



These do *not* look like white noise!

ACF of model residuals



There is significant autocorrelation at lag = 1

Model with autocorrelated errors

We can expand the stochastic part of the model to have autocorrelated errors

$$
\begin{aligned}
y_t &= \alpha + \beta x_t + e_t \\
e_t &= \phi e_{t-1} + w_t
\end{aligned}
$$

with $w_t \sim \text{N}(0, \sigma^2)$

Model with autocorrelated errors

We can expand the stochastic part of the model to have autocorrelated errors

$$
\begin{aligned}
y_t &= \alpha + \beta x_t + e_t \\
e_t &= \phi e_{t-1} + w_t
\end{aligned}
$$

with $w_t \sim \mathrm{N}(0, \sigma^2)$

We can write this model as our standard state-space model

State-space model

Observation equation

$$
\begin{aligned}
y_t &= \alpha + \beta x_t + e_t \\
&= e_t + \alpha + \beta x_t \\
&\Downarrow \\
y_t &= x_t + a + D d_t + v_t
\end{aligned}
$$

with

$$x_t = e_t, a = \alpha, D = \beta, d_t = x_t, v_t = 0$$

State-space model

State equation

$$e_t = \phi e_{t-1} + w_t$$
$$\Downarrow$$
$$x_t = B x_t + w_t$$

with

$x_t = e_t$ and $B = \phi$

State-space model

Full form

$$
\begin{aligned}
y_t &= \alpha + \beta x_t + e_t \\
e_t &= \phi e_{t-1} + w_t \\
&\Downarrow \\
y_t &= a + D d_t + x_t \\
x_t &= B x_t + w_t
\end{aligned}
$$

State-space model

Observation model in `MARSS()`

$$y_t = a + Dd_t + x_t$$
$$\Downarrow$$
$$y_t = Zx_t + a + Dd_t + v_t$$

```
y = data          ## [1 x T] matrix of data
a = matrix("a")   ## intercept
D = matrix("D")   ## slope
d = covariate     ## [1 x T] matrix of measured covariate
Z = matrix(1)     ## no multiplier on x
R = matrix(0)     ## v_t ~ N(0,R); want v_t = 0 for all t
```

State-space model

State model in `MARSS()`

$$x_t = Bx_t + w_t$$
$$\Downarrow$$
$$x_t = Bx_t + u + Cc_t + w_t$$

```
B = matrix("b")  ## AR(1) coefficient for model errors
Q = matrix("q")  ## w_t ~ N(0,Q); var for model errors
u = matrix(0)    ## u = 0
C = matrix(0)    ## C = 0
c = matrix(0)    ## c_t = 0 for all t
```

MORE RANDOM EFFECTS

Expanding the random effect

Recall our simple model

$$y_t = \underbrace{\mu}_{\text{fixed}} + \underbrace{e_t}_{\text{random}}$$

Expanding the random effect

We can expand the random portion

$$y_t = \underbrace{\mu}_{\text{fixed}} + \underbrace{f_t + e_t}_{\text{random}}$$

$$e_t \sim \text{N}(0, \sigma)$$
$$f_t \sim \text{N}(f_{t-1}, \gamma)$$

Expanding the random effect

We can expand the random portion

$$y_t = \underbrace{\mu}_{\text{fixed}} + \underbrace{f_t + e_t}_{\text{random}}$$

$$e_t \sim \mathrm{N}(0, \sigma)$$
$$f_t \sim \mathrm{N}(f_{t-1}, \gamma)$$

This is simply a random walk observed with error

Random walk observed with error

$$y_t = \mu + f_t + e_t \text{ with } e_t \sim \mathrm{N}(0, \sigma)$$
$$f_t = f_{t-1} + w_t \text{ with } w_t \sim \mathrm{N}(0, \gamma)$$
$$\Downarrow$$
$$y_t = a + x_t + v_t \text{ with } v_t \sim \mathrm{N}(0, R)$$
$$x_t = x_{t-1} + w_t \text{ with } w_t \sim \mathrm{N}(0, Q)$$

Expanding fixed & random effects

We can expand the fixed portion

$$y_t = \underbrace{\alpha + \beta x_t}_{\text{fixed}} + \underbrace{f_t + e_t}_{\text{random}}$$

$$e_t \sim \text{N}(0, \sigma)$$
$$f_t \sim \text{N}(f_{t-1}, \gamma)$$

Fixed & random effects

In familiar state-space form

$$
\begin{gathered}
y_t = \alpha + \beta x_t + f_t + e_t \text{ with } e_t \sim \text{N}(0, \sigma) \\
f_t = f_{t-1} + w_t \text{ with } w_t \sim \text{N}(0, \gamma) \\
\Downarrow \\
y_t = a + D d_t + x_t + v_t \text{ with } v_t \sim \text{N}(0, R) \\
x_t = x_{t-1} + w_t \text{ with } w_t \sim \text{N}(0, Q)
\end{gathered}
$$

MULTIPLE TIME SERIES

Simple model for 2+ time series

Random walk observed with error

$$
\begin{aligned}
y_{i,t} &= x_{i,t} + a_i + v_{i,t} \\
x_{i,t} &= x_{i,t-1} + w_{i,t}
\end{aligned}
$$

with

$v_{i,t} \sim \text{N}(0, R)$

$w_{i,t} \sim \text{N}(0, Q)$

Random walk observed with error

$$
\begin{aligned}
y_{1,t} &= x_{1,t} + a_1 + v_{1,t} \\
y_{2,t} &= x_{2,t} + a_2 + v_{2,t} \\
&\vdots \\
y_{n,t} &= x_{n,t} + a_2 + v_{n,t}
\end{aligned}
$$

$$
\begin{aligned}
x_{1,t} &= x_{1,t-1} + w_{1,t} \\
x_{2,t} &= x_{2,t-1} + w_{2,t} \\
&\vdots \\
x_{n,t} &= x_{n,t-1} + w_{n,t}
\end{aligned}
$$

Random walk observed with error

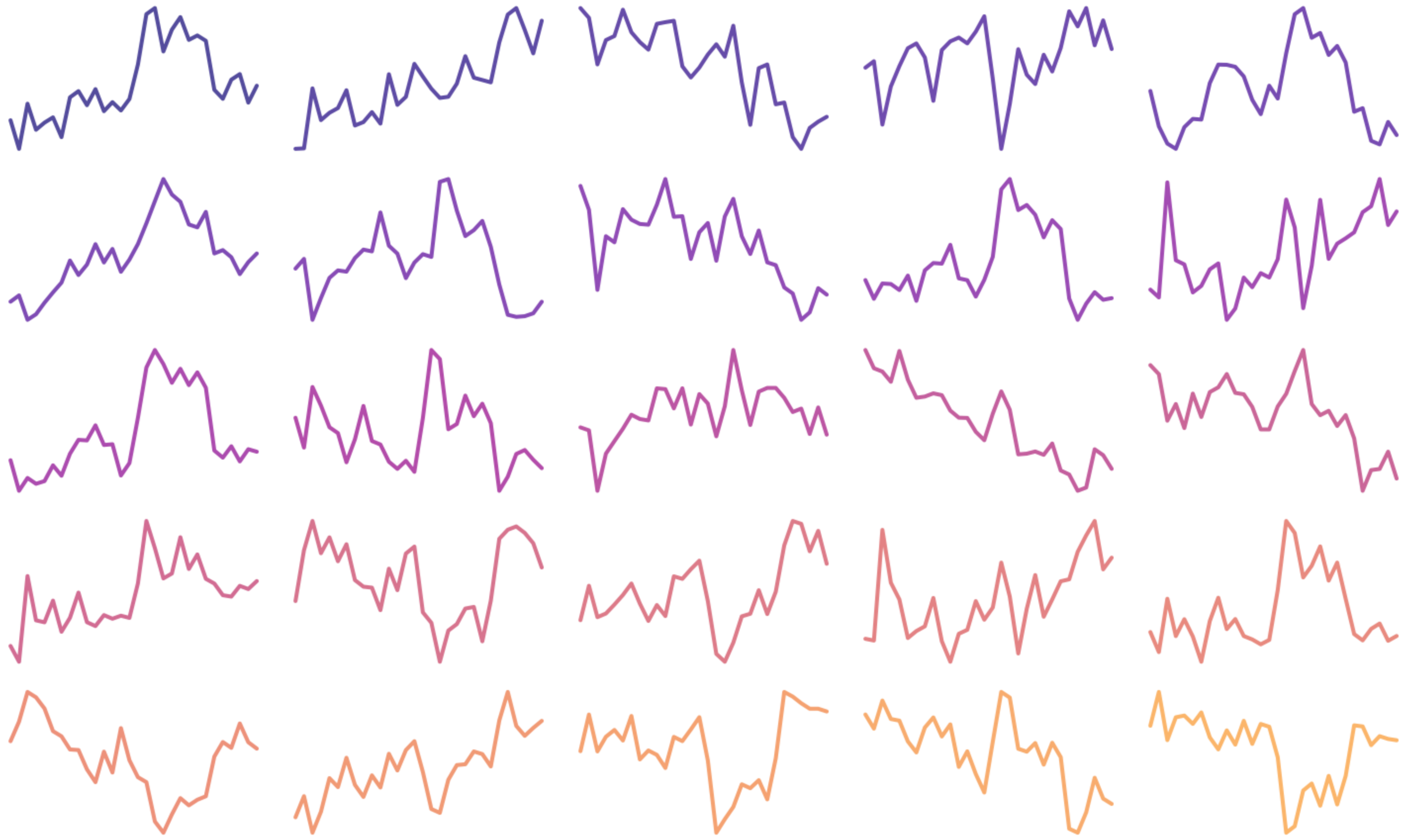
In matrix form

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

with

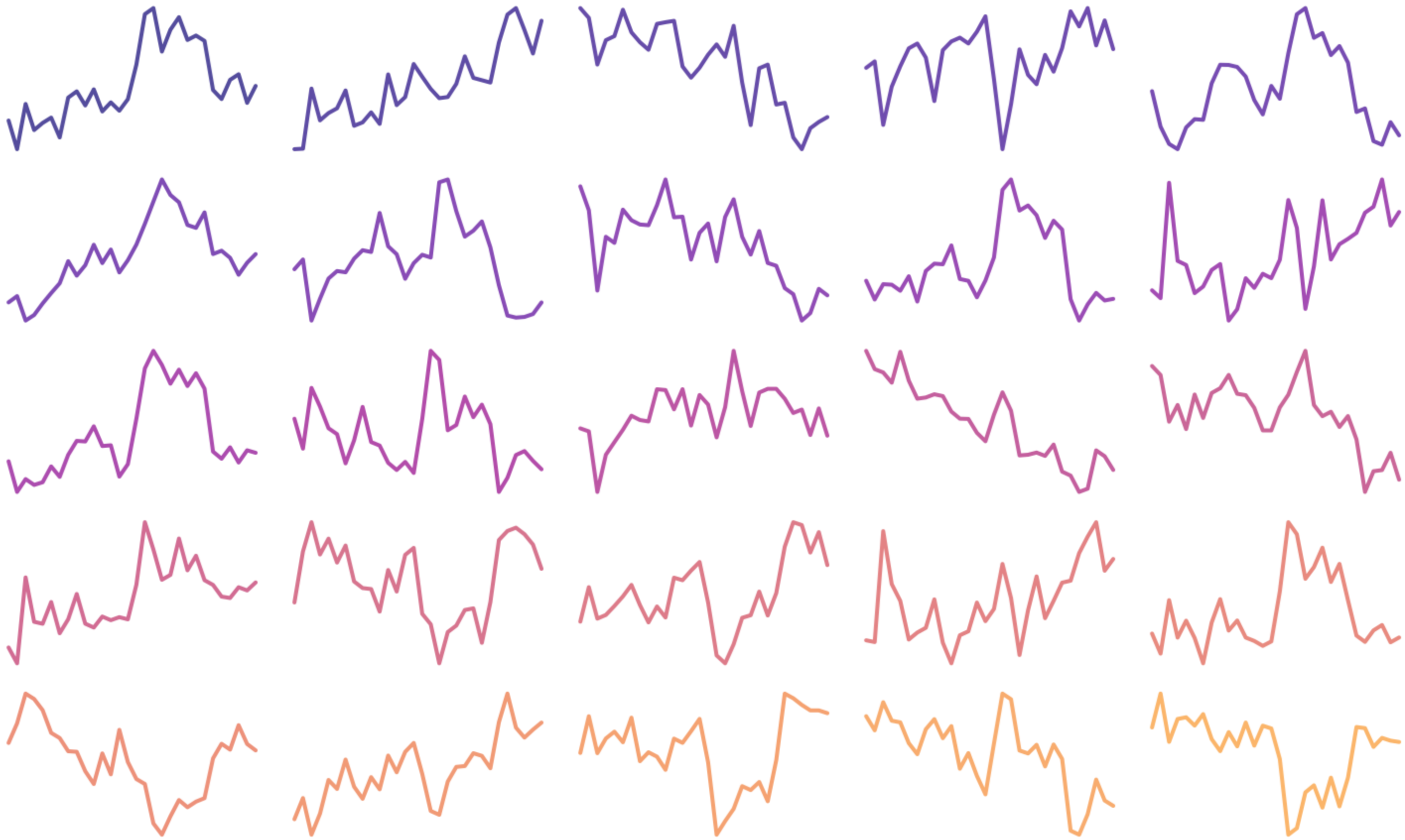
$\mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$

$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$



Environmental time series

We often observe covariance among environmental time series, especially for those collected close to one another in space



Are there some common patterns here?

Common patterns in time series

State-space model

Ex: population structure

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

We can make (test) assumptions by specifying different forms for $\mathbf{Z}$

State-space model

Ex: Harbor seal population structure

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}_t = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_t + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t$$

$$\begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_t = \begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_{t-1} + \begin{bmatrix} w_{JF} \\ w_N \\ w_S \end{bmatrix}_t$$

Finding common patterns

What if our observations were instead a mixture of 2+ states?

For example, we sampled haul-outs located between several breeding sites

Mixtures of states

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \end{bmatrix}_t = \begin{bmatrix} 0.8 & 0.2 & 0 \\ 0.2 & 0.7 & 0.1 \\ 0 & 0.9 & 0.1 \\ 0 & 0.3 & 0.7 \\ 0 & 0.1 & 0.9 \end{bmatrix} \times \begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_t + \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ v_4 \\ v_5 \end{bmatrix}_t$$

$$\begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_t = \begin{bmatrix} x_{JF} \\ x_N \\ x_S \end{bmatrix}_{t-1} + \begin{bmatrix} w_{JF} \\ w_N \\ w_S \end{bmatrix}_t$$

Finding common patterns

What if our observations were a mixture of states, but we didn't know how many or the weightings?

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

What are the dimensions of $\mathbf{Z}$?

What are the elements within $\mathbf{Z}$?

Dynamic Factor Analysis (DFA)

DFA is a *dimension reduction* technique, which models n observed time series as a function of m hidden states (patterns), where $n \gg m$

Dynamic Factor Analysis (DFA)

State-space form

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

data: $\mathbf{y}_t$ is $n \times 1$

loadings: $\mathbf{Z}$ is $n \times m$ with $n > m$

states: $\mathbf{x}_t$ is $m \times 1$

Dimension reduction

Principal Components Analysis (PCA)

Goal is to reduce some large number of correlated variates into a few uncorrelated factors

Principal Components Analysis (PCA)

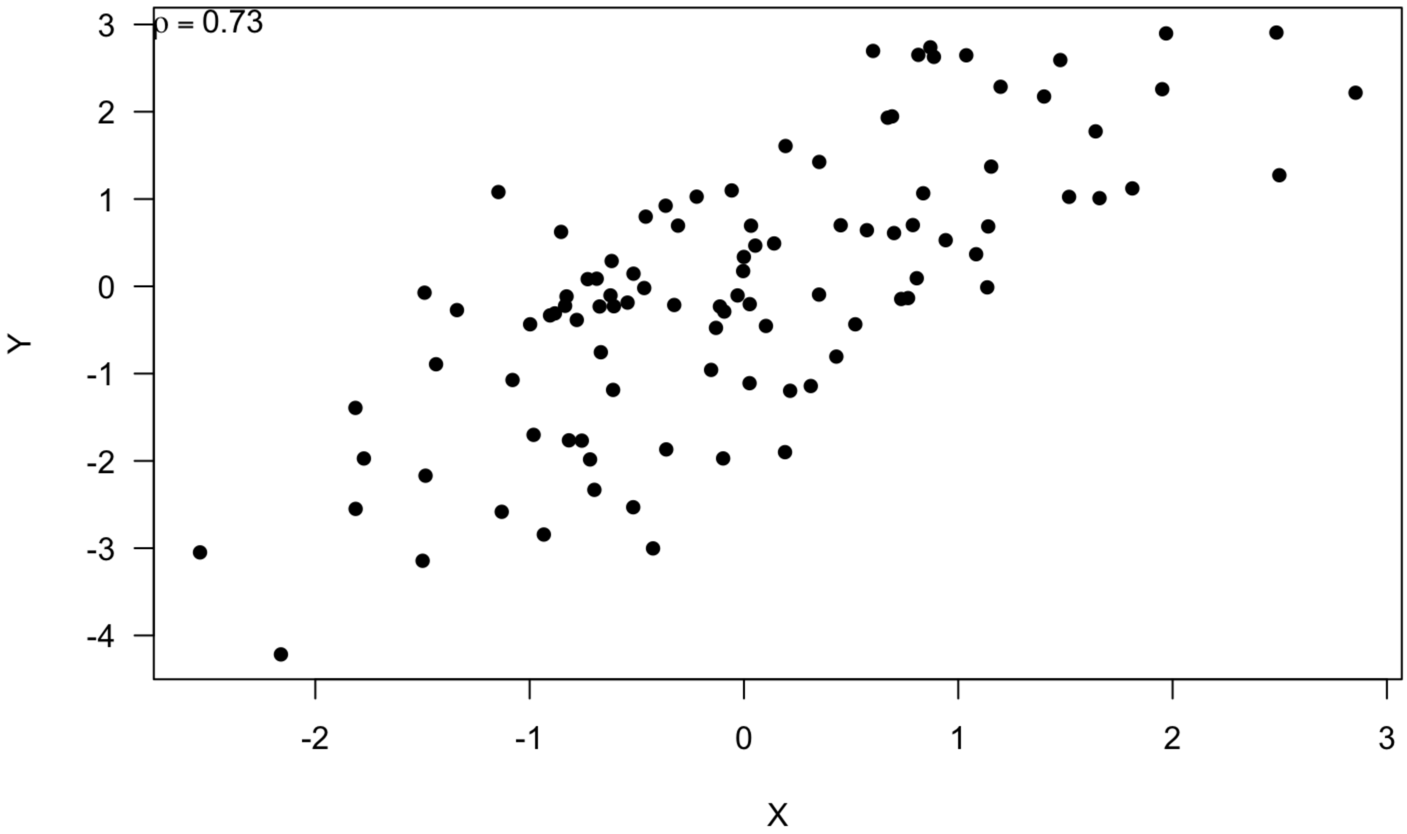
Calculating the principal components requires us to estimate the covariance of the data

$$\text{PC} = \text{eigenvectors}(\text{cov}(\mathbf{y}))$$

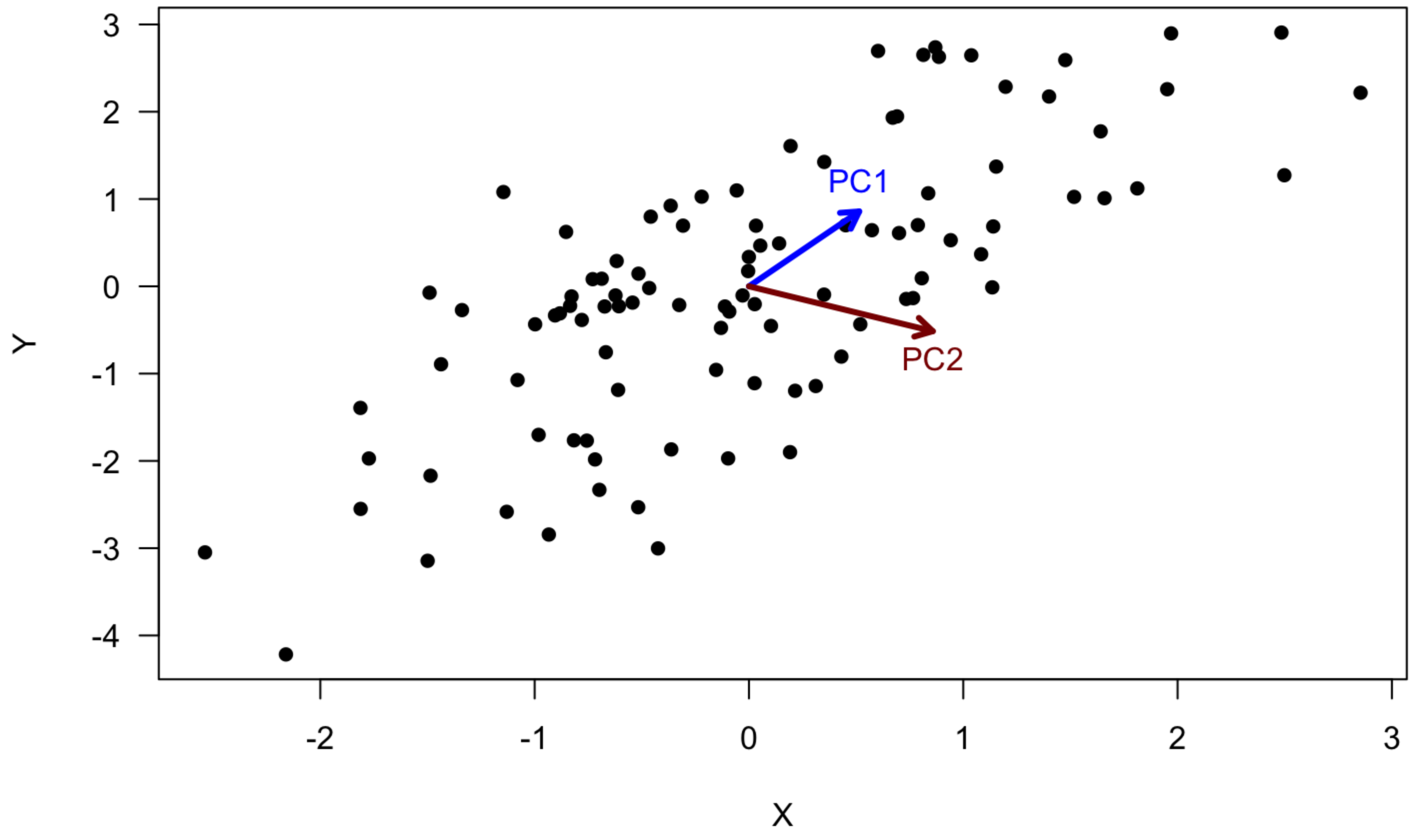
There will be n principal components (eigenvectors) for an $n \times T$ matrix $\mathbf{y}$

We reduce the dimension by selecting a subset of the components that explain much of the variance (eg, the first 2)

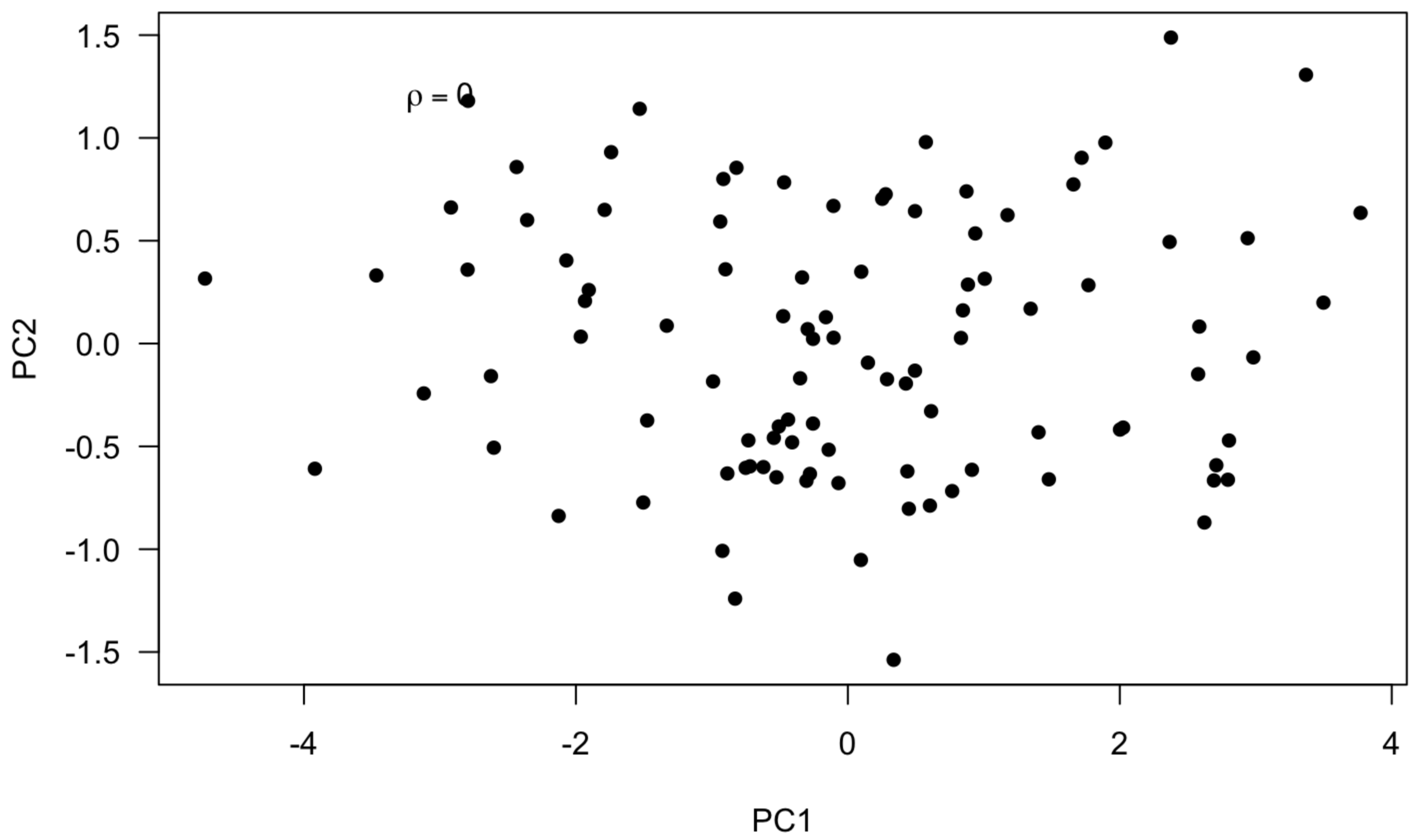
Principal Components Analysis (PCA)



Principal Components Analysis (PCA)



Principal Components Analysis (PCA)



Relationship between PCA & DFA

We need to estimate the covariance in the data $\mathbf{y}$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t, \text{ with } \mathbf{v}_t \sim \text{MVN}(\mathbf{0}, \mathbf{R})$$

so

$$\text{cov}(\mathbf{y}_t) = \mathbf{Z}\text{cov}(\mathbf{x}_t)\mathbf{Z}^\top + \mathbf{R}$$

In PCA, we require $\mathbf{R}$ to be diagonal, but not so in DFA

Principal Components Analysis (PCA)

Forms for $\mathbf{R}$ with $n = 4$

$$\mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma & 0 & 0 & 0 \\ 0 & \sigma & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & \sigma \end{bmatrix} \text{ or } \mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}$$

Dynamic Factor Analysis (DFA)

Forms for $\mathbf{R}$ with $n = 4$

$$\mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma & 0 & 0 & 0 \\ 0 & \sigma & 0 & 0 \\ 0 & 0 & \sigma & 0 \\ 0 & 0 & 0 & \sigma \end{bmatrix} \text{ or } \mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & 0 \\ 0 & 0 & \sigma_3 & 0 \\ 0 & 0 & 0 & \sigma_4 \end{bmatrix}$$

$$\mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma & \gamma & \gamma & \gamma \\ \gamma & \sigma & \gamma & \gamma \\ \gamma & \gamma & \sigma & \gamma \\ \gamma & \gamma & \gamma & \sigma \end{bmatrix} \text{ or } \mathbf{R} \stackrel{?}{=} \begin{bmatrix} \sigma_1 & 0 & 0 & 0 \\ 0 & \sigma_2 & 0 & \gamma_{2,4} \\ 0 & 0 & \sigma_3 & 0 \\ 0 & \gamma_{2,4} & 0 & \sigma_4 \end{bmatrix}$$

Dynamic Factor Analysis (DFA)

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

What form should we use for $\mathbf{Z}$?

$$\mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \\ z_5 \end{bmatrix} \text{ or } \mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \\ z_{1,3} & z_{2,3} \\ z_{1,4} & z_{2,4} \\ z_{1,5} & z_{2,5} \end{bmatrix} \text{ or } \mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_{1,1} & z_{2,1} & z_{3,1} \\ z_{1,2} & z_{2,2} & z_{3,2} \\ z_{1,3} & z_{2,3} & z_{3,3} \\ z_{1,4} & z_{2,4} & z_{3,4} \\ z_{1,5} & z_{2,5} & z_{3,5} \end{bmatrix}$$

Dynamic Factor Analysis (DFA)

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\ \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t \end{aligned}$$

What form should we use for $\mathbf{Z}$?

$$\mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \\ z_5 \end{bmatrix} \text{ or } \mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \\ z_{1,3} & z_{2,3} \\ \vdots & \vdots \\ z_{1,n} & z_{2,n} \end{bmatrix} \text{ or } \mathbf{Z} \stackrel{?}{=} \begin{bmatrix} z_{1,1} & z_{2,1} & z_{3,1} \\ z_{1,2} & z_{2,2} & z_{3,2} \\ z_{1,3} & z_{2,3} & z_{3,3} \\ \vdots & \vdots & \vdots \\ z_{1,n} & z_{2,n} & z_{3,n} \end{bmatrix}$$

We’ll use model selection criteria to choose (eg, AICc)

Fitting DFA models

Unless $\mathbf{Z}$ is unconstrained in some manner, there are an infinite number of combinations of $\mathbf{Z}$ and $\mathbf{x}$ that will equal $\mathbf{y}$

Therefore we need to impose some constraints on the model

Constraints on DFA models

1) The offset $\mathbf{a}$

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

$$
\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}
$$

Constraints on DFA models

1) The offset $\mathbf{a}$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ \vdots \\ a_n \end{bmatrix}$$

We will set the first m elements of $\mathbf{a}$ to 0

Constraints on DFA models

1) The offset $\mathbf{a}$

For example, if $n = 5$ and $m = 2$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \Rightarrow \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix}$$

Constraints on DFA models

1) The offset $\mathbf{a}$

For example, if $n = 5$ and $m = 2$

$$\mathbf{a} = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \Rightarrow \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ a_3 \\ a_4 \\ a_5 \end{bmatrix} \Rightarrow \mathbf{a} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

Note, however, that this causes problems for the EM algorithm so we will often de-mean the data and set $a_i = 0$ for all i

Constraints on DFA models

2) The loadings $\mathbf{Z}$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{Z} = \begin{bmatrix} z_{1,1} & z_{2,1} & \cdots & z_{m,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{m,2} \\ z_{1,3} & z_{2,3} & \cdots & z_{m,3} \\ \vdots & \vdots & \ddots & z_{m,4} \\ z_{1,n} & z_{2,n} & \cdots & z_{m,n} \end{bmatrix}$$

Constraints on DFA models

2) The loadings $\mathbf{Z}$

$$\mathbf{y}_t = \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t$$
$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{w}_t$$

$$\mathbf{Z} = \begin{bmatrix} z_{1,1} & z_{2,1} & \cdots & z_{m,1} \\ z_{1,2} & z_{2,2} & \cdots & z_{m,2} \\ z_{1,3} & z_{2,3} & \cdots & z_{m,3} \\ \vdots & \vdots & \ddots & z_{m,4} \\ z_{1,n} & z_{2,n} & \cdots & z_{m,n} \end{bmatrix}$$

We will set the upper right triangle of $\mathbf{Z}$ to 0

Constraints on DFA models

2) The loadings $\mathbf{Z}$

For example, if $n = 5$ and $m = 3$

$$\mathbf{Z} = \begin{bmatrix} z_{1,1} & 0 & 0 \\ z_{1,2} & z_{2,2} & 0 \\ z_{1,3} & z_{2,3} & z_{3,3} \\ z_{1,4} & z_{2,3} & z_{3,4} \\ z_{1,5} & z_{2,5} & z_{3,5} \end{bmatrix}$$

For the first $m - 1$ rows of $\mathbf{Z}$, $z_{i,j} = 0$ if $j > i$

Constraints on DFA models

2) The loadings $\mathbf{Z}$

An additional constraint is necessary in a Bayesian context

$$\mathbf{Z} = \begin{bmatrix} \underline{z_{1,1}} & 0 & 0 \\ z_{1,2} & \underline{z_{2,2}} & 0 \\ z_{1,3} & z_{2,3} & \underline{z_{3,3}} \\ z_{1,4} & z_{2,3} & z_{3,4} \\ z_{1,5} & z_{2,5} & z_{3,5} \end{bmatrix}$$

Diagonal of $\mathbf{Z}$ is positive: $z_{i,j} > 0$ if $i = j$

Constraints on DFA models

3) The state variance $\mathbf{Q}$

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

$$
\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})
$$

Constraints on DFA models

3) The state variance $\mathbf{Q}$

$$\begin{aligned}\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{v}_t \\ \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t\end{aligned}$$

$$\mathbf{w}_t \sim \text{MVN}(\mathbf{0}, \mathbf{Q})$$

We will set $\mathbf{Q}$ equal to the Identity matrix $\mathbf{I}$

Constraints on DFA models

3) The state variance $\mathbf{Q}$

For example, if $m = 4$

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

This allows our random walks to have a *lot* of flexibility

Dynamic Factor Analysis (DFA)

Including p covariates

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \underline{\mathbf{D}\mathbf{d}_t} + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

$\mathbf{d}_t$ is a $p \times 1$ vector of covariates at time t

$\mathbf{D}$ is an $n \times p$ matrix of covariate effects

Dynamic Factor Analysis (DFA)

Form for $\mathbf{D}$

$$
\begin{aligned}
\mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \underline{\mathbf{D}}\mathbf{d}_t + \mathbf{v}_t \\
\mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t
\end{aligned}
$$

Careful thought must be given *a priori* as to the form for $\mathbf{D}$

Should the effect(s) vary by site, species, etc?

Dynamic Factor Analysis (DFA)

Form for $\mathbf{D}$

For example, given 2 covariates, Temp and Salinity

$$
\mathbf{D} = \underbrace{\begin{bmatrix} d_{\text{Temp}} & d_{\text{Salinity}} \\ d_{\text{Temp}} & d_{\text{Salinity}} \\ \vdots & \vdots \\ d_{\text{Temp}} & d_{\text{Salinity}} \end{bmatrix}}_{\text{effects same by site/species}} \quad \text{or} \quad \mathbf{D} = \underbrace{\begin{bmatrix} d_{\text{Temp},1} & d_{\text{Salinity},1} \\ d_{\text{Temp},2} & d_{\text{Salinity},2} \\ \vdots & \vdots \\ d_{\text{Temp},n} & d_{\text{Salinity},n} \end{bmatrix}}_{\text{effects differ by site/species}}
$$

A note on model selection

Earlier we saw that we could use model selection criteria to help us choose among the different forms for $\mathbf{Z}$

However, caution must be given when comparing models with and without covariates, and varying numbers of states

A note on model selection

Think about the DFA model form

$$\mathbf{y}_t = \mathbf{Z}\underline{\mathbf{x}_t} + \mathbf{a} + \mathbf{D}\underline{\mathbf{d}_t} + \mathbf{v}_t$$

$\mathbf{x}_t$ are *undetermined* random walks

$\mathbf{d}_t$ are *predetermined* covariates

An example with 3 times series

Model 1 has 2 trends and no covariates

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \\ z_{1,3} & z_{2,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t$$

Model 2 has 1 trend and 1 covariate

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \begin{bmatrix} x \end{bmatrix}_t + \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \begin{bmatrix} d \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t$$

An example with 3 times series

Model 1 has 2 trends and no covariates

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_{1,1} & z_{2,1} \\ z_{1,2} & z_{2,2} \\ z_{1,3} & z_{2,3} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t$$

Model 2 has 1 trend and 1 covariate

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix}_t = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix} \begin{bmatrix} x \end{bmatrix}_t + \begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} \begin{bmatrix} d \end{bmatrix}_t + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}_t$$

Unless $\mathbf{d}$ is *highly correlated* with $\mathbf{y}$, Model 1 will be favored

A note on model selection

For models with covariates

- fit the *most complex model you can envision* based on all of your possible covariates and random factors (states)
- keep the covariates fixed and choose the number of trends (states) using AICc
- keep the covariates & states fixed and choose the form for $\mathbf{R}$
- sort out the covariates while keeping the states & $\mathbf{R}$ fixed

Interpreting DFA results

Recall that we had to constrain the form of $\mathbf{Z}$ to fit the model

$$
\mathbf{Z} = \begin{bmatrix}
z_{1,1} & 0 & \dots & 0 \\
z_{1,2} & z_{2,2} & \ddots & 0 \\
\vdots & \vdots & \ddots & 0 \\
\vdots & \vdots & \vdots & z_{m,m} \\
\vdots & \vdots & \vdots & \vdots \\
z_{1,n} & z_{2,n} & z_{3,n} & z_{m,n}
\end{bmatrix}
$$

So, the 1st common factor is determined by the 1st variate, the 2nd common factor by the first two variates, etc.

Interpreting DFA results

To help with this, we can use a *basis rotation* to maximize the loadings on a few factors

If $\mathbf{H}$ is an $m \times m$ non-singular matrix, these 2 DFA models are equivalent

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t \\ \mathbf{x}_t &= \mathbf{x}_{t-1} + \mathbf{w}_t \end{aligned}$$

$$\begin{aligned} \mathbf{y}_t &= \mathbf{Z}\mathbf{H}^{-1}\mathbf{x}_t + \mathbf{a} + \mathbf{D}\mathbf{d}_t + \mathbf{v}_t \\ \mathbf{H}\mathbf{x}_t &= \mathbf{H}\mathbf{x}_{t-1} + \mathbf{H}\mathbf{w}_t \end{aligned}$$

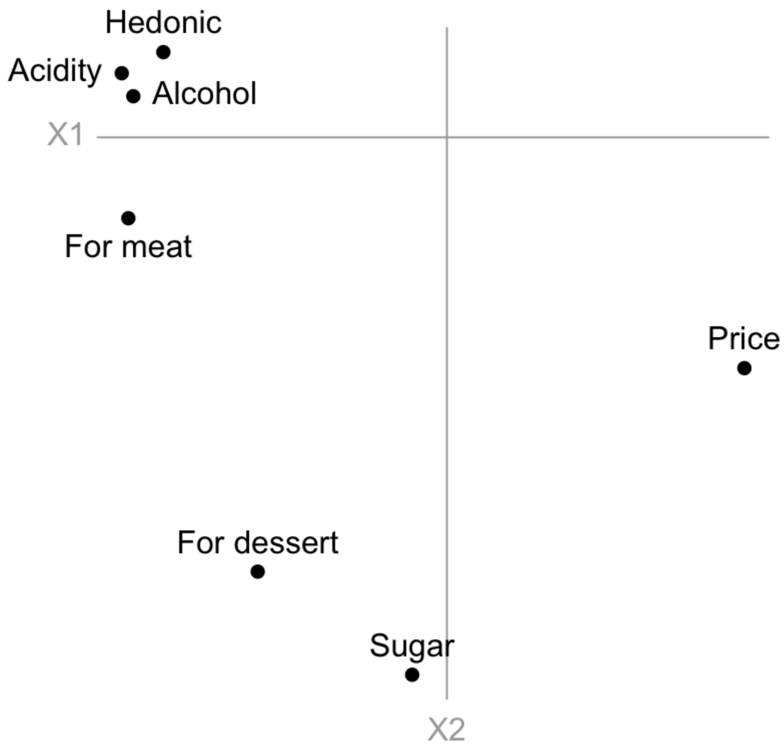
How should we choose $\mathbf{H}$?

Basis rotation

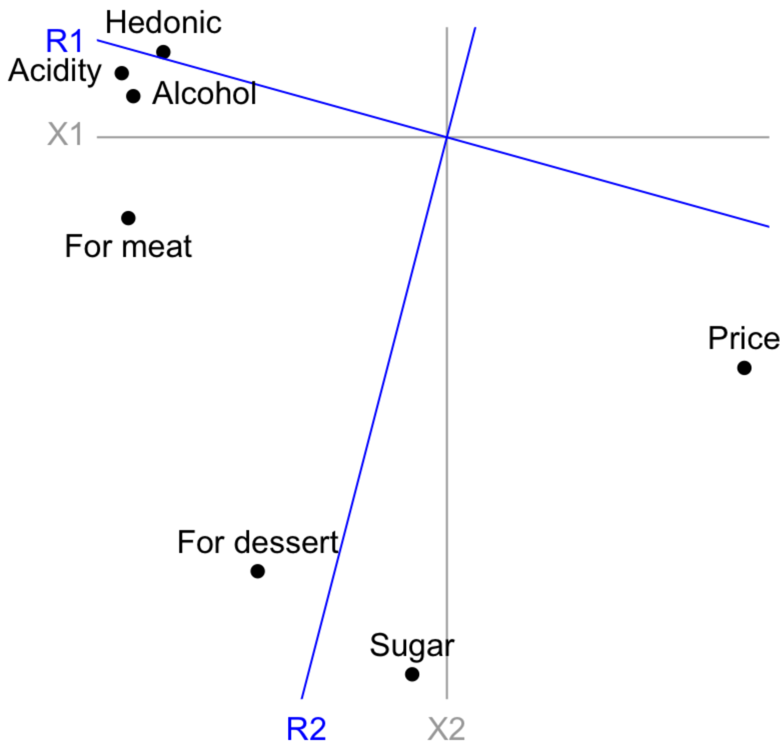
Varimax

A *varimax* rotation will maximize the variance of the loadings in $\mathbf{Z}$ along a few of the factors

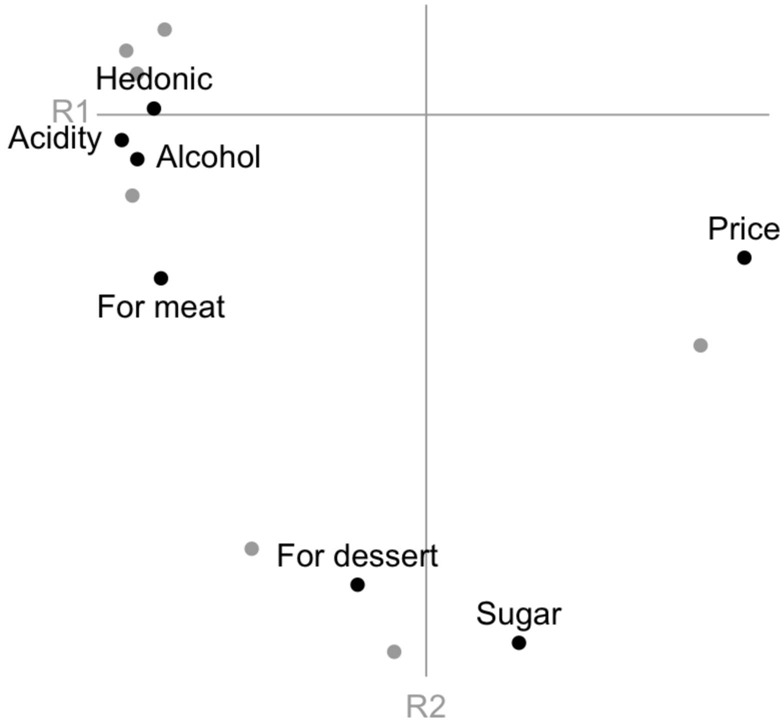
PCA of 5 wines with 8 attributes



Rotated loadings



Rotated loadings



Topics for today

Deterministic vs stochastic elements

Regression with autocorrelated errors

Regression with temporal random effects

Dynamic Factor Analysis (DFA)

- Forms of covariance matrix
- Constraints for model fitting
- Interpretation of results