Introduction to Time Series Analysis. Lecture 7.

Peter Bartlett

Last lecture:

1. ARMA(p,q) models: stationarity, causality, invertibility
2. The linear process representation of ARMA processes: ψ.
3. Autocovariance of an ARMA process.
4. Homogeneous linear difference equations.

Introduction to Time Series Analysis. Lecture 7.

Peter Bartlett

1. Review: ARMA(p,q) models and their properties
2. Review: Autocovariance of an ARMA process.
3. Homogeneous linear difference equations.

Forecasting

1. Linear prediction.
2. Projection in Hilbert space.

Review: Autoregressive moving average models

An **ARMA(p,q) process** $\{X_t\}$ is a stationary process that satisfies

$$\phi(B)X_t = \theta(B)W_t,$$

where ϕ, θ are degree p, q polynomials and $\{W_t\} \sim WN(0, \sigma^2)$.

We'll insist that the polynomials ϕ and θ have no common factors.

Review: Properties of ARMA(p,q) models

Theorem: If ϕ and θ have no common factors, a (unique) *stationary* solution $\{X_t\}$ to $\phi(B)X_t = \theta(B)W_t$
exists iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \;\Rightarrow\; |z| \neq 1.$$

This ARMA(p,q) process is *causal* iff

$$\phi(z) = 1 - \phi_1 z - \cdots - \phi_p z^p = 0 \;\Rightarrow\; |z| > 1.$$

It is *invertible* iff

$$\theta(z) = 1 + \theta_1 z + \cdots + \theta_q z^q = 0. \;\Rightarrow\; |z| > 1.$$

Review: Properties of ARMA(p,q) models

$$\phi(B)X_t = \theta(B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t$$

so $\qquad \theta(B) = \psi(B)\phi(B)$

$$\Leftrightarrow \vdots$$

$$\begin{aligned} \Leftrightarrow \qquad 1 &= \psi_0, \\ \theta_1 &= \psi_1 - \phi_1\psi_0, \\ \theta_2 &= \psi_2 - \phi_1\psi_1 - \cdots - \phi_2\psi_0, \\ &\vdots \end{aligned}$$

We need to solve the linear difference equation $\theta_j = \phi(B)\psi_j$, with $\theta_0 = 1$, $\theta_j = 0$ for $j < 0$, $j > q$.

Review: Autocovariance functions of ARMA processes

$$\phi(B)X_t = \theta(B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t,$$

$$\gamma(h) - \phi_1\gamma(h-1) - \phi(2)\gamma(h-2) - \cdots = \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}.$$

We need to solve the homogeneous linear difference equation $\phi(B)\gamma(h) = 0$ ($h > q$), with initial conditions

$$\gamma(h) - \phi_1\gamma(h-1) - \cdots - \phi_p\gamma(h-p) = \sigma_w^2 \sum_{j=h}^{q} \theta_j \psi_{j-h}, \qquad h = 0, \ldots, q.$$

Introduction to Time Series Analysis. Lecture 7.

1. Review: ARMA(p,q) models and their properties
2. Review: Autocovariance of an ARMA process.
3. Homogeneous linear difference equations.

Forecasting

1. Linear prediction.
2. Projection in Hilbert space.

Homogeneous linear diff eqns with constant coefficients

$$
\begin{aligned}
a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} &= 0 \\
\Leftrightarrow \qquad \left(a_0 + a_1 B + \cdots + a_k B^k\right) x_t &= 0 \\
\Leftrightarrow \qquad a(B) x_t &= 0 \\
\text{auxiliary equation:} \qquad a_0 + a_1 z + \cdots + a_k z^k &= 0 \\
\Leftrightarrow \qquad (z - z_1)(z - z_2) \cdots (z - z_k) &= 0
\end{aligned}
$$

where $z_1, z_2, \ldots, z_k \in \mathbb{C}$ are the roots of this *characteristic polynomial*.

Thus,

$$
a(B) x_t = 0 \qquad \Leftrightarrow \qquad (B - z_1)(B - z_2) \cdots (B - z_k) x_t = 0.
$$

Homogeneous linear diff eqns with constant coefficients

$$a(B)x_t = 0 \qquad \Leftrightarrow \qquad (B - z_1)(B - z_2)\cdots(B - z_k)x_t = 0.$$

So any $\{x_t\}$ satisfying $(B - z_i)x_t = 0$ for some i also satisfies $a(B)x_t = 0$.

Three cases:

1. The z_i are real and distinct.
2. The z_i are complex and distinct.
3. Some z_i are repeated.

Homogeneous linear diff eqns with constant coefficients

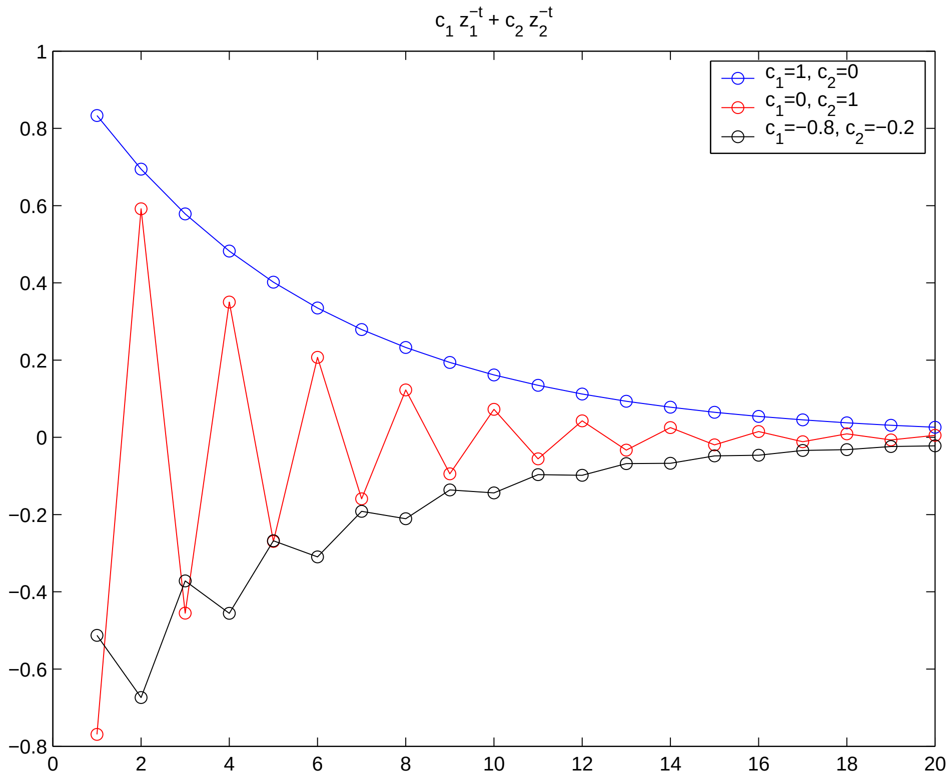
1. The z_i are real and distinct.

$$
\begin{aligned}
& & a(B)x_t = 0 \\
&\Leftrightarrow & (B - z_1)(B - z_2)\cdots(B - z_k)x_t = 0 \\
&\Leftrightarrow & x_t \text{ is a linear combination of solutions to} \\
& & (B - z_1)x_t = 0,\ (B - z_2)x_t = 0, \ldots,\ (B - z_k)x_t = 0 \\
&\Leftrightarrow & x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t},
\end{aligned}
$$

for some constants $c_1, \ldots, c_k$.

Homogeneous linear diff eqns with constant coefficients

1. The z_i are real and distinct. e.g., $z_1 = 1.2$, $z_2 = -1.3$



Reminder: Complex exponentials

$$a + ib = re^{i\theta} = r(\cos\theta + i\sin\theta),$$

$$\text{where} \quad r = |a + ib| = \sqrt{a^2 + b^2}$$

$$\theta = \tan^{-1}\left(\frac{b}{a}\right) \in (-\pi, \pi].$$

Thus,

$$r_1 e^{i\theta_1} r_2 e^{i\theta_2} = (r_1 r_2) e^{i(\theta_1 + \theta_2)},$$

$$z\bar{z} = |z|^2.$$

Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct.

As before, $$a(B)x_t = 0$$

$$\Leftrightarrow \qquad x_t = c_1 z_1^{-t} + c_2 z_2^{-t} + \cdots + c_k z_k^{-t}.$$

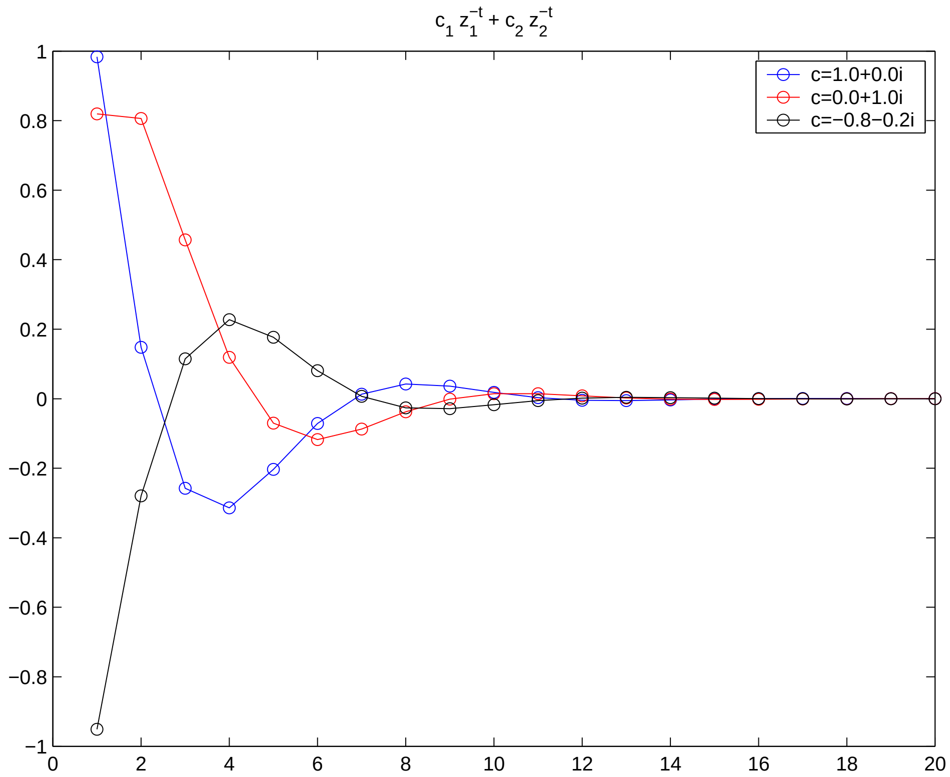
If $z_1 \notin \mathbb{R}$, since $a_1, \ldots, a_k$ are real, we must have the complex conjugate root, $z_j = \bar{z_1}$. And for x_t to be real, we must have $c_j = \bar{c_1}$. For example:

$$\begin{aligned} x_t &= c\, z_1^{-t} + \bar{c}\, \bar{z_1}^{-t} \\ &= r\, e^{i\theta} |z_1|^{-t} e^{-i\omega t} + r\, e^{-i\theta} |z_1|^{-t} e^{i\omega t} \\ &= r|z_1|^{-t} \left(e^{i(\theta - \omega t)} + e^{-i(\theta - \omega t)} \right) \\ &= 2r|z_1|^{-t} \cos(\omega t - \theta) \end{aligned}$$

where $z_1 = |z_1| e^{i\omega}$ and $c = re^{i\theta}$.

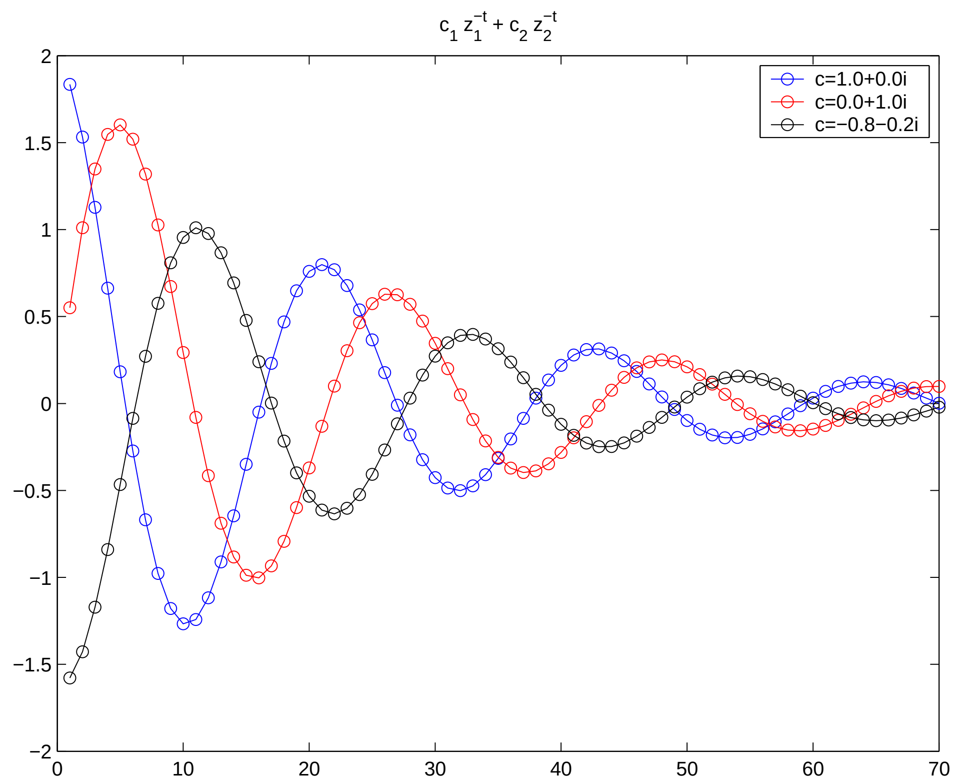
Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct. e.g., $z_1 = 1.2 + i$, $z_2 = 1.2 - i$



Homogeneous linear diff eqns with constant coefficients

2. The z_i are complex and distinct. e.g., $z_1 = 1 + 0.1i$, $z_2 = 1 - 0.1i$



Homogeneous linear diff eqns with constant coefficients

3. Some z_i are repeated.

$$
\begin{aligned}
& a(B)x_t = 0 \\
\Leftrightarrow \quad & (B - z_1)(B - z_2)\cdots(B - z_k)x_t = 0.
\end{aligned}
$$

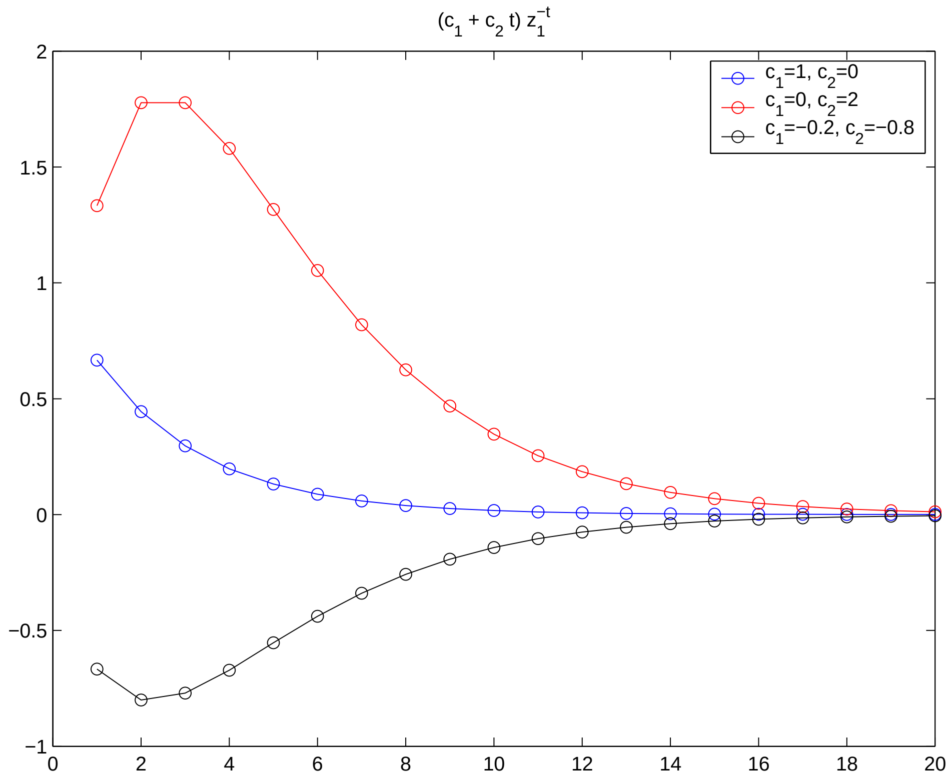
$$
\begin{aligned}
\text{If } z_1 = z_2, \quad & (B - z_1)(B - z_2)x_t = 0 \\
\Leftrightarrow \quad & (B - z_1)^2 x_t = 0.
\end{aligned}
$$

We can check that $(c_1 + c_2 t)z_1^{-t}$ is a solution...

More generally, $(B - z_1)^m x_t = 0$ has the solution $\left(c_1 + c_2 t + \cdots + c_{m-1} t^{m-1}\right) z_1^{-t}$.

Homogeneous linear diff eqns with constant coefficients

3. Some z_i are repeated. e.g., $z_1 = z_2 = 1.5$.



Solving linear diff eqns with constant coefficients

$$\textbf{Solve:} \qquad a_0 x_t + a_1 x_{t-1} + \cdots + a_k x_{t-k} = 0,$$

$$\text{with initial conditions} \quad x_1, \ldots, x_k.$$

$$\text{Auxiliary equation in } z \in \mathbb{C}\text{:} \qquad a_0 + a_1 z + \cdots + a_k z^k = 0$$

$$\Leftrightarrow \qquad (z - z_1)^{m_1} (z - z_2)^{m_2} \cdots (z - z_l)^{m_l} = 0,$$

where $z_1, z_2, \ldots, z_l \in \mathbb{C}$ are the roots of the characteristic polynomial, and z_i occurs with multiplicity m_i.

Solutions: $c_1(t) z_1^{-t} + c_2(t) z_2^{-t} + \cdots + c_l(t) z_l^{-t}$,

where $c_i(t)$ is a polynomial in t of degree $m_i - 1$.

We determine the coefficients of the $c_i(t)$ using the initial conditions (which might be linear constraints on the initial values $x_1, \ldots, x_k$).

Autocovariance functions of ARMA processes: Example

$$(1+0.25B^2)X_t = (1+0.2B)W_t, \qquad \Leftrightarrow \qquad X_t = \psi(B)W_t,$$

$$\psi_j = \left(1, \frac{1}{5}, -\frac{1}{4}, -\frac{1}{20}, \frac{1}{16}, \frac{1}{80}, -\frac{1}{64}, -\frac{1}{320}, \ldots\right).$$

$$\gamma(h) - \phi_1\gamma(h-1) - \phi_2\gamma(h-2) = \sigma_w^2 \sum_{j=0}^{q-h} \theta_{h+j}\psi_j$$

$$\Leftrightarrow \gamma(h) + 0.25\gamma(h-2) = \begin{cases} \sigma_w^2\left(\psi_0 + 0.2\psi_1\right) & \text{if } h = 0, \\ 0.2\sigma_w^2\psi_0 & \text{if } h = 1, \\ 0 & \text{otherwise.} \end{cases}$$

Autocovariance functions of ARMA processes: Example

We have the homogeneous linear difference equation

$$\gamma(h) + 0.25\gamma(h-2) = 0$$

for $h \geq 2$, with initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2 \,(1 + 1/25)$$
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

Autocovariance functions of ARMA processes: Example

Homogeneous lin. diff. eqn:

$$\gamma(h) + 0.25\gamma(h-2) = 0.$$

The characteristic polynomial is

$$1 + 0.25z^2 = \frac{1}{4}\left(4 + z^2\right) = \frac{1}{4}(z - 2i)(z + 2i),$$

which has roots at $z_1 = 2e^{i\pi/2}$, $\bar{z_1} = 2e^{-i\pi/2}$.

The solution is of the form

$$\gamma(h) = cz_1^{-h} + \bar{c}\bar{z_1}^{-h}.$$

Autocovariance functions of ARMA processes: Example

$z_1 = 2e^{i\pi/2}, \bar{z_1} = 2e^{-i\pi/2}, c = |c|e^{i\theta}.$

We have

$$\begin{aligned}\gamma(h) &= c z_1^{-h} + \bar{c}\bar{z_1}^{-h} \\ &= 2^{-h}\left(|c|e^{i(\theta - h\pi/2)} + |c|e^{i(-\theta + h\pi/2)}\right) \\ &= c_1 2^{-h}\cos\left(\frac{h\pi}{2} - \theta\right).\end{aligned}$$

And we determine c_1, θ from the initial conditions

$$\gamma(0) + 0.25\gamma(-2) = \sigma_w^2\,(1 + 1/25)$$

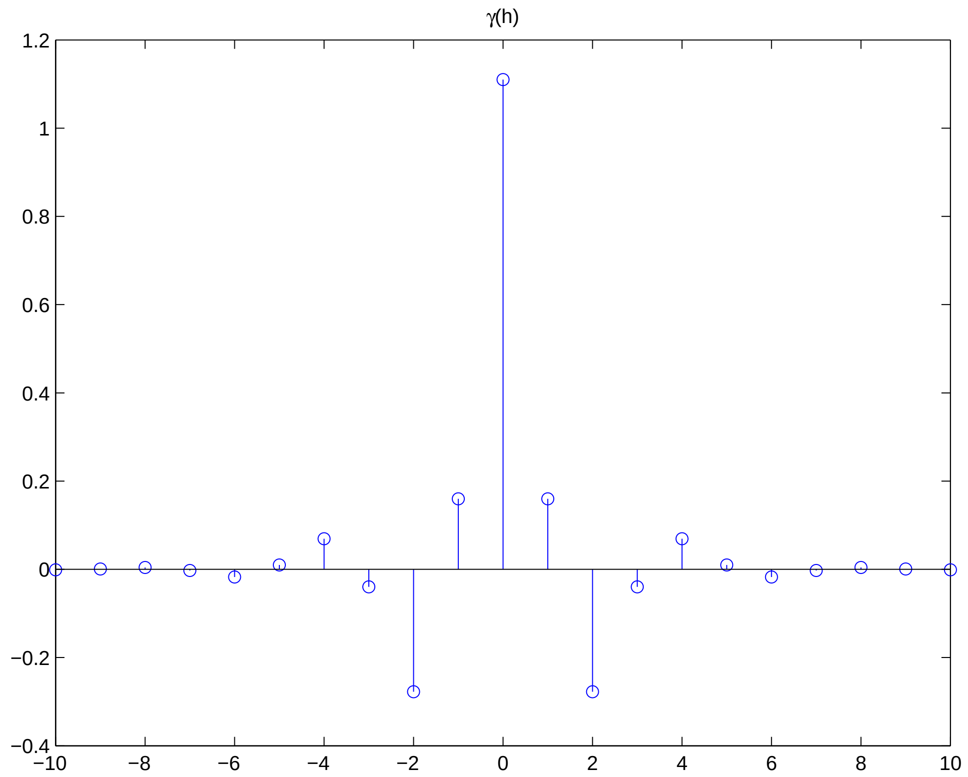
$$\gamma(1) + 0.25\gamma(-1) = \sigma_w^2/5.$$

Autocovariance functions of ARMA processes: Example

We determine c_1, θ from the initial conditions:

$$\begin{aligned}
\text{We plug} \qquad \gamma(0) &= c_1 \cos(\theta) \\
\gamma(1) &= \frac{c_1}{2} \sin(\theta) \\
\gamma(2) &= -\frac{c_1}{4} \cos(\theta) \\
\text{into} \qquad \gamma(0) + 0.25\gamma(2) &= \sigma_w^2 \,(1 + 1/25) \\
1.25\gamma(1) &= \sigma_w^2/5.
\end{aligned}$$

Autocovariance functions of ARMA processes: Example



Introduction to Time Series Analysis. Lecture 7.

1. Review: ARMA(p,q) models and their properties
2. Review: Autocovariance of an ARMA process.
3. Homogeneous linear difference equations.

Forecasting

1. Linear prediction.
2. Projection in Hilbert space.

Review: least squares linear prediction

Consider a **linear predictor** of X_{n+h} given $X_n = x_n$:

$$f(x_n) = \alpha_0 + \alpha_1 x_n.$$

For a stationary time series $\{X_t\}$, the best linear predictor is $f^*(x_n) = (1 - \rho(h))\mu + \rho(h)x_n$:

$$\begin{aligned} \mathrm{E}\left(X_{n+h} - (\alpha_0 + \alpha_1 X_n)\right)^2 &\geq \mathrm{E}\left(X_{n+h} - f^*(X_n)\right)^2 \\ &= \sigma^2(1 - \rho(h)^2). \end{aligned}$$

Linear prediction

Given $X_1, X_2, \ldots, X_n$, the best linear predictor

$$X_{n+m}^n = \alpha_0 + \sum_{i=1}^n \alpha_i X_i$$

of X_{n+m} satisfies the **prediction equations**

$$\mathrm{E}\left(X_{n+m} - X_{n+m}^n\right) = 0$$

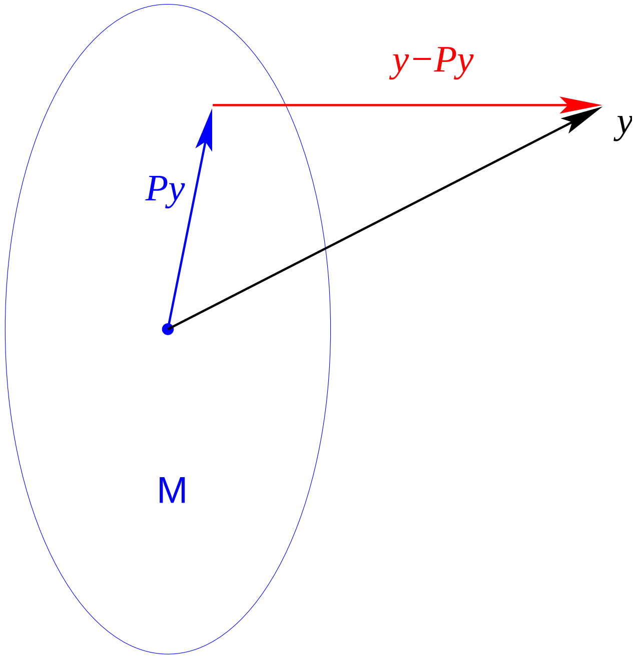
$$\mathrm{E}\left[\left(X_{n+m} - X_{n+m}^n\right) X_i\right] = 0 \qquad \text{for } i = 1, \ldots, n.$$

This is a special case of the *projection theorem*.

Projection Theorem

If $\mathcal{H}$ is a Hilbert space,
$\mathcal{M}$ is a closed linear subspace of $\mathcal{H}$,
and $y \in \mathcal{H}$,
then there is a point $Py \in \mathcal{M}$
(the **projection of** y **on** $\mathcal{M}$)
satisfying

1. $\|Py - y\| \leq \|w - y\|$ for $w \in \mathcal{M}$,
2. $\|Py - y\| < \|w - y\|$ for $w \in \mathcal{M}$, $w \neq y$
3. $\langle y - Py, w \rangle = 0$ for $w \in \mathcal{M}$.



Hilbert spaces

Hilbert space $=_{\text{complete}}$ inner product space:

Inner product space: vector space, with inner product $\langle a, b\rangle$:

- $\langle a, b\rangle = \langle b, a\rangle$,
- $\langle \alpha_1 a_1 + \alpha_2 a_2, b\rangle = \alpha_1\langle a_1, b\rangle + \alpha_2\langle a_2, b\rangle$,
- $\langle a, a\rangle = 0 \Leftrightarrow a = 0$.

Norm: $\|a\|^2 = \langle a, a\rangle$.

complete = limits of Cauchy sequences are in the space

Examples:

1. $\mathbb{R}^n$, with Euclidean inner product, $\langle x, y\rangle = \sum_i x_i y_i$.
2. $\{$random variables X: $\mathrm{E}X^2 < \infty\}$,
with inner product $\langle X, Y\rangle = \mathrm{E}(XY)$. (Strictly, equivalence classes of a.s. equal r.v.s)

Projection theorem

Example: **Linear regression**

Given $y = (y_1, y_2, \ldots, y_n)' \in \mathbb{R}^n$, and $Z = (z_1, \ldots, z_q) \in \mathbb{R}^{n \times q}$, choose $\beta = (\beta_1, \ldots, \beta_q)' \in \mathbb{R}^q$ to minimize $\|y - Z\beta\|^2$.

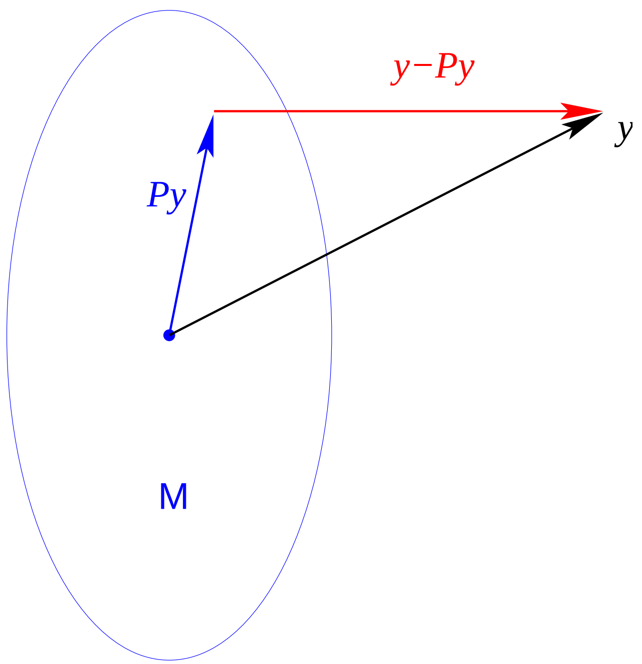
Here, $\mathcal{H} = \mathbb{R}^n$, with $\langle a, b \rangle = \sum_i a_i b_i$, and
$\mathcal{M} = \{Z\beta : \beta \in \mathbb{R}^q\} = \overline{\text{sp}}\,\{z_1, \ldots, z_q\}$.

Projection theorem

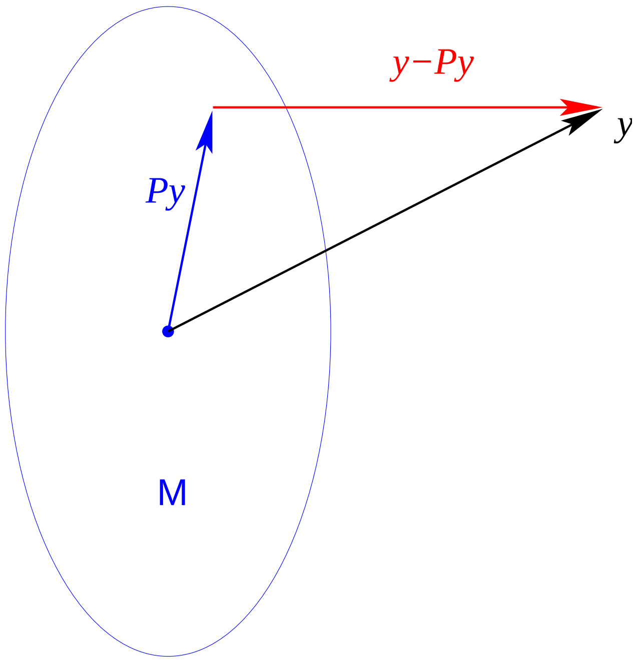
If $\mathcal{H}$ is a Hilbert space,
$\mathcal{M}$ is a closed subspace of $\mathcal{H}$,
and $y \in \mathcal{H}$,
then there is a point $Py \in \mathcal{M}$
(the **projection of** y **on** $\mathcal{M}$)
satisfying

1. $\|Py - y\| \leq \|w - y\|$
2. $\langle y - Py, w \rangle = 0$

for $w \in \mathcal{M}$.



Projection theorem



$$\langle y - Py, w \rangle = 0$$

$$\Leftrightarrow \quad \langle y - Z\hat{\beta}, z_i \rangle = 0, \quad \forall i$$

$$\Leftrightarrow \quad Z'Z\hat{\beta} = Z'y$$

$$\Leftrightarrow \quad \hat{\beta} = (Z'Z)^{-1}Z'y$$

"normal equations."

Projection theorem

Example: **Linear prediction**

Given $1, X_1, X_2, \ldots, X_n \in \left\{\text{r.v.s } X : \mathrm{E}X^2 < \infty\right\}$,

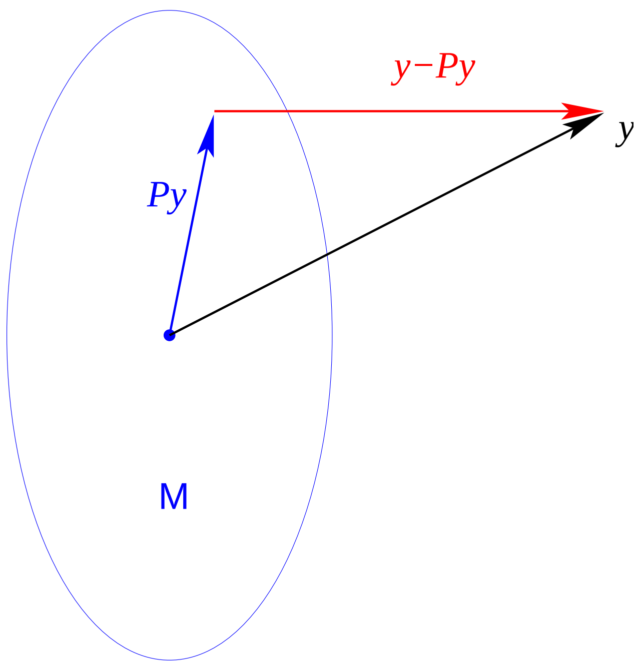
choose $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$
so that $Z = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i$ minimizes $\mathrm{E}(X_{n+m} - Z)^2$.

Here, $\langle X, Y \rangle = \mathrm{E}(XY)$,
$\mathcal{M} = \{Z = \alpha_0 + \sum_{i=1}^{n} \alpha_i X_i : \alpha_i \in \mathbb{R}\} = \bar{\text{sp}}\{1, X_1, \ldots, X_n\}$, and
$y = X_{n+m}$.

Projection theorem

If $\mathcal{H}$ is a Hilbert space,
$\mathcal{M}$ is a closed subspace of $\mathcal{H}$,
and $y \in \mathcal{H}$,
then there is a point $Py \in \mathcal{M}$
(the **projection of y on $\mathcal{M}$**)
satisfying
1. $\|Py - y\| \leq \|w - y\|$
2. $\langle y - Py, w \rangle = 0$

for $w \in \mathcal{M}$.



Projection theorem: Linear prediction

Let X^n_{n+m} denote the best linear predictor:

$$\|X^n_{n+m} - X_{n+m}\|^2 \leq \|Z - X_{n+m}\|^2 \quad \text{for all } Z \in \mathcal{M}.$$

The projection theorem implies the orthogonality

$$\begin{aligned} & \langle X^n_{n+m} - X_{n+m}, Z \rangle = 0 \quad \text{for all } Z \in \mathcal{M} \\ \Leftrightarrow \quad & \langle X^n_{n+m} - X_{n+m}, Z \rangle = 0 \quad \text{for all } Z \in \{1, X_1, \ldots, X_n\} \end{aligned}$$

$$\Leftrightarrow \quad \begin{aligned} & \mathrm{E}\left(X^n_{n+m} - X_{n+m}\right) = 0 \\ & \mathrm{E}\left[\left(X^n_{n+m} - X_{n+m}\right) X_i\right] = 0 \end{aligned}$$

That is, the *prediction errors ($X^n_{n+m} - X_{n+m}$) are uncorrelated with the prediction variables ($1, X_1, \ldots, X_n$).*

Linear prediction

Error $(X^n_{n+m} - X_{n+m})$ is uncorrelated with the prediction variable 1:

$$
\begin{aligned}
& & \mathrm{E}\left(X^n_{n+m} - X_{n+m}\right) = 0 \\
\Leftrightarrow & & \mathrm{E}\left(\alpha_0 + \sum_i \alpha_i X_i - X_{n+m}\right) = 0 \\
\Leftrightarrow & & \mu\left(1 - \sum_i \alpha_i\right) = \alpha_0.
\end{aligned}
$$

Linear prediction

... $$\mu\left(1-\sum_i \alpha_i\right)=\alpha_0.$$

Substituting for α_0 in

$$X^n_{n+m}=\alpha_0+\sum_i \alpha_i X_i,$$

we get $$X^n_{n+m}=\mu+\sum_i \alpha_i\left(X_i-\mu\right).$$

So we can subtract μ from all variables:

$$X^n_{n+m}-\mu=\sum_i \alpha_i\left(X_i-\mu\right).$$

Thus, for forecasting, we can assume $\mu=0$. So we'll ignore α_0.

One-step-ahead linear prediction

Write $X_{n+1}^n = \phi_{n1} X_n + \phi_{n2} X_{n-1} + \cdots + \phi_{nn} X_1$

Prediction equations: $\mathrm{E}\left((X_{n+1}^n - X_{n+1}) X_i \right) = 0, \text{ for } i = 1, \ldots, n$

$$\Leftrightarrow \quad \sum_{j=1}^{n} \phi_{nj} \mathrm{E}\left(X_{n+1-j} X_i \right) = \mathrm{E}(X_{n+1} X_i)$$

$$\Leftrightarrow \quad \sum_{j=1}^{n} \phi_{nj} \gamma(i-j) = \gamma(i)$$

$$\Leftrightarrow \quad \Gamma_n \phi_n = \gamma_n,$$

One-step-ahead linear prediction

Prediction equations: $\Gamma_n \phi_n = \gamma_n$.

$$\Gamma_n = \begin{bmatrix} \gamma(0) & \gamma(1) & \cdots & \gamma(n-1) \\ \gamma(1) & \gamma(0) & & \gamma(n-2) \\ \vdots & & \ddots & \vdots \\ \gamma(n-1) & \gamma(n-2) & \cdots & \gamma(0) \end{bmatrix},$$

$$\phi_n = (\phi_{n1}, \phi_{n2}, \ldots, \phi_{nn})', \quad \gamma_n = (\gamma(1), \gamma(2), \ldots, \gamma(n))'.$$

Mean squared error of one-step-ahead linear prediction

$$
\begin{aligned}
P_{n+1}^n &= \mathrm{E}\left(X_{n+1} - X_{n+1}^n\right)^2 \\
&= \mathrm{E}\left(\left(X_{n+1} - X_{n+1}^n\right)\left(X_{n+1} - X_{n+1}^n\right)\right) \\
&= \mathrm{E}\left(X_{n+1}\left(X_{n+1} - X_{n+1}^n\right)\right) \\
&= \gamma(0) - \mathrm{E}\left(\phi_n' X X_{n+1}\right) \\
&= \gamma(0) - \gamma_n' \Gamma_n^{-1} \gamma_n,
\end{aligned}
$$

where $X = (X_n, X_{n-1}, \ldots, X_1)'$.

Mean squared error of one-step-ahead linear prediction

Variance is reduced:

$$\begin{aligned}
P_{n+1}^{n} &= \mathrm{E}\left(X_{n+1} - X_{n+1}^{n}\right)^{2} \\
&= \gamma(0) - \gamma_n' \Gamma_n^{-1} \gamma_n \\
&= \mathrm{Var}(X_{n+1}) - \mathrm{Cov}(X_{n+1}, X)\mathrm{Cov}(X, X)^{-1}\mathrm{Cov}(X, X_{n+1}) \\
&= \mathrm{E}\left(X_{n+1} - 0\right)^{2} - \mathrm{Cov}(X_{n+1}, X)\mathrm{Cov}(X, X)^{-1}\mathrm{Cov}(X, X_{n+1}),
\end{aligned}$$

where $X = (X_n, X_{n-1}, \ldots, X_1)'$.

Introduction to Time Series Analysis. Lecture 7.

Peter Bartlett

1. Review: ARMA(p,q) models and their properties
2. Review: Autocovariance of an ARMA process.
3. Homogeneous linear difference equations.

Forecasting

1. Linear prediction.
2. Projection in Hilbert space.