Introduction to Time Series Analysis. Lecture 12.

Peter Bartlett

1. Review: Time series modelling and forecasting
2. Parameter estimation
3. Maximum likelihood estimator
4. Yule-Walker estimation

Review (Lecture 1): Time series modelling and forecasting

1. Plot the time series.

   Look for trends, seasonal components, step changes, outliers.

2. Transform data so that residuals are **stationary**.

   (a) Remove trend and seasonal components.

   (b) Differencing.

   (c) Nonlinear transformations (log, $\sqrt{\cdot}$).

3. Fit model to residuals.

4. Forecast time series by forecasting residuals and inverting any transformations.

Review: Time series modelling and forecasting

Stationary time series models: ARMA(p,q).

- $p = 0$: MA(q),
- $q = 0$: AR(p).

We have seen that any causal, invertible linear process has:
an MA(∞) representation (from causality), and
an AR(∞) representation (from invertibility).

Real data cannot be *exactly* modelled using a finite number of parameters.

We choose p, q to give a simple but accurate model.

Review: Time series modelling and forecasting

How do we use data to decide on p, q?

1. Use sample ACF/PACF to make preliminary choices of model order.
2. Estimate parameters for each of these choices.
3. Compare predictive accuracy/complexity of each (using, e.g., AIC).

NB: We need to compute parameter estimates for several different model orders.

Thus, recursive algorithms for parameter estimation are important.

We'll see that some of these are identical to the recursive algorithms for forecasting.

Review: Time series modelling and forecasting

Model:	ACF:	PACF:
AR(p)	decays	zero for $h > p$
MA(q)	zero for $h > q$	decays
ARMA(p,q)	decays	decays

Introduction to Time Series Analysis. Lecture 12.

1. Review: Time series modelling and forecasting
2. Parameter estimation
3. Maximum likelihood estimator
4. Yule-Walker estimation

Parameter estimation

We want to estimate the parameters of an ARMA(p,q) model.
We will assume (for now) that:
1. The model order (p and q) is known, and
2. The data has zero mean.

If (2) is not a reasonable assumption, we can subtract the sample mean $\bar{y}$, fit a zero-mean ARMA model,

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

to the mean-corrected time series $X_t = Y_t - \bar{y}$, and then use $X_t + \bar{y}$ as the model for Y_t.

Parameter estimation: Maximum likelihood estimator

One approach:

Assume that $\{X_t\}$ is Gaussian, that is, $\phi(B)X_t = \theta(B)W_t$, where W_t is i.i.d. Gaussian.

Choose ϕ_i, θ_j to maximize the *likelihood*:

$$L(\phi, \theta, \sigma^2) = f(X_1, \ldots, X_n),$$

where f is the joint (Gaussian) density for the given ARMA model.

(c.f. choosing the parameters that maximize the probability of the data.)

Parameter estimation: Maximum likelihood estimator

Advantages of MLE:

Efficient (low variance estimates).

Often the Gaussian assumption is reasonable.

Even if $\{X_t\}$ is not Gaussian, the asymptotic distribution of the estimates $(\hat{\phi}, \hat{\theta}, \hat{\sigma}^2)$ is the same as the Gaussian case.

Disadvantages of MLE:

Difficult optimization problem.

Need to choose a good starting point (often use other estimators for this).

Preliminary parameter estimates

Yule-Walker for AR(p): Regress X_t onto $X_{t-1},\ldots,X_{t-p}$.

Durbin-Levinson algorithm with γ replaced by $\hat{\gamma}$.

Yule-Walker for ARMA(p,q): Method of moments. Not efficient.

Innovations algorithm for MA(q): with γ replaced by $\hat{\gamma}$.

Hannan-Rissanen algorithm for ARMA(p,q):

1. Estimate high-order AR.
2. Use to estimate (unobserved) noise W_t.
3. Regress X_t onto $X_{t-1},\ldots,X_{t-p},\hat{W}_{t-1},\ldots,\hat{W}_{t-q}$.
4. Regress again with improved estimates of W_t.

Yule-Walker estimation

For a causal AR(p) model $\phi(B)X_t = W_t$, we have

$$\mathrm{E}\left(X_{t-i}\left(X_t - \sum_{j=1}^{p}\phi_j X_{t-j}\right)\right) = \mathrm{E}(X_{t-i}W_t) \quad \text{for } i = 0, \ldots, p$$

$$\Leftrightarrow \qquad \gamma(0) - \phi'\gamma_p = \sigma^2 \quad \text{and}$$

$$\gamma_p - \Gamma_p\phi = 0,$$

where $\phi = (\phi_1, \ldots, \phi_p)'$, and we've used the causal representation

$$X_t = W_t + \sum_{j=1}^{\infty}\psi_j W_{t-j}.$$

Yule-Walker estimation

Method of moments: We choose parameters for which the moments are equal to the empirical moments.

In this case, we choose ϕ so that $\gamma = \hat{\gamma}$.

Yule-Walker equations for $\hat{\phi}$:
$$\begin{cases} \hat{\Gamma}_p \hat{\phi} = \hat{\gamma}_p, \\ \hat{\sigma}^2 = \hat{\gamma}(0) - \hat{\phi}' \hat{\gamma}_p. \end{cases}$$

These are the forecasting equations.
We can use the Durbin-Levinson algorithm.

Some facts about Yule-Walker estimation

- If $\hat{\gamma}(0) > 0$, then $\hat{\Gamma}_m$ is nonsingular.
- In that case, $\hat{\phi} = \hat{\Gamma}_p^{-1}\gamma_p$ defines the causal model

$$X_t - \hat{\phi}_1 X_{t-1} - \cdots - \hat{\phi}_p X_{t-p} = W_t, \quad \{W_t\} \sim WN(0, \hat{\sigma}^2).$$

- If $\{X_t\}$ is an AR(p) process,

$$\hat{\phi} \sim AN\left(\phi, \frac{\sigma^2}{n}\Gamma_p^{-1}\right), \qquad \hat{\sigma}^2 \xrightarrow{P} \sigma^2.$$

$$\hat{\phi}_{hh} \sim AN\left(0, \frac{1}{n}\right) \quad \text{for } h > p.$$

Thus, we can use the sample PACF to test for AR order, and we can calculate approximate confidence intervals for the parameters ϕ.

Yule-Walker estimation: Confidence intervals

If $\{X_t\}$ is an AR(p) process, and n is large,

- $\sqrt{n}(\hat{\phi}_p - \phi_p)$ is approximately $N(0, \hat{\sigma}^2 \hat{\Gamma}_p^{-1})$,
- with probability $\approx 1 - \alpha$, ϕ_p is in the ellipsoid

$$\left\{ \phi \in \mathbb{R}^p : \left(\hat{\phi}_p - \phi \right)' \hat{\Gamma}_p \left(\hat{\phi}_p - \phi \right) \leq \frac{\hat{\sigma}^2}{n} \chi^2_{1-\alpha}(p) \right\},$$

where $\chi^2_{1-\alpha}(p)$ is the $(1-\alpha)$ quantile of the chi-squared with p degrees of freedom.

- with probability $\approx 1 - \alpha$, ϕ_{pj} is in the interval

$$\hat{\phi}_{pj} \pm \Phi_{1-\alpha/2} \frac{\hat{\sigma}}{\sqrt{n}} \left(\hat{\Gamma}_p^{-1} \right)_{jj}^{1/2},$$

where $\Phi_{1-\alpha/2}$ is the $1 - \alpha/2$ quantile of the standard normal.

Introduction to Time Series Analysis. Lecture 12.

1. Review: Time series modelling and forecasting
2. Parameter estimation
3. Maximum likelihood estimator
4. Yule-Walker estimation