1 The AR(1) Model

The AR(1) model

An autoregressive process is one in which the current value depends on its own past values and an unpredictable shock. The simplest autoregressive model is

$$y_t = m + \phi y_{t-1} + \varepsilon_t.$$

Because the model uses one lag of y_t, it is called an autoregressive process of order 1, or AR(1).

The shock ε_t is white noise. In particular,

$$E(\varepsilon_t) = 0,$$

and shocks are uncorrelated across time.

For a stationary AR(1),

$$|\phi| < 1.$$

Mean and the mean-centered form

A stationary AR(1) has a constant unconditional mean μ. Taking expectations of

$$y_t = m + \phi y_{t-1} + \varepsilon_t$$

gives

$$\mu = m + \phi\mu.$$

Therefore,

$$\mu = \frac{m}{1-\phi}, \qquad m = (1-\phi)\mu.$$

Using this relationship, we can write the same AR(1) as

$$y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t.$$

This form is useful because it shows directly how deviations from the mean persist over time. If $0 < \phi < 1$, only a fraction ϕ of today's deviation from the mean is expected to remain next period.

Forecasting with an AR(1)

Suppose T is the last observed period. From

$$y_{T+1} = m + \phi y_T + \varepsilon_{T+1},$$

the one-step-ahead forecast is

$$\widehat{y}_{T+1|T} = m + \phi y_T,$$

because the future shock has conditional mean zero.

Equivalently, using the mean-centered form,

$$\widehat{y}_{T+1|T} = \mu + \phi(y_T - \mu).$$

For forecasts more than one period ahead, use the **chain rule of forecasting**: use the previous forecast in place of the future value that has not yet been observed. For example,

$$\widehat{y}_{T+2|T} = m + \phi\widehat{y}_{T+1|T}.$$

More generally,

$$\widehat{y}_{T+h|T} = \mu + \phi^h(y_T - \mu).$$

Mean reversion

For a stationary AR(1), $|\phi| < 1$, so

$$\phi^h \to 0 \qquad \text{as } h \to \infty.$$

Therefore,

$$\widehat{y}_{T+h|T} \to \mu.$$

Thus, long-horizon forecasts of a stationary AR(1) revert toward the unconditional mean. When $0 < \phi < 1$, a larger ϕ means greater persistence and slower mean reversion.

Forecast errors

The one-step-ahead forecast error is

$$e_{T+1|T} = y_{T+1} - \widehat{y}_{T+1|T} = \varepsilon_{T+1}.$$

The two-step-ahead forecast error is

$$e_{T+2|T} = \varepsilon_{T+2} + \phi\varepsilon_{T+1}.$$

Thus, longer-horizon forecasts depend on more future shocks. Forecast uncertainty therefore increases with the forecast horizon.

Exercise 1 Forecasting with an AR(1).

Suppose inflation follows

$$y_t - 4 = 0.5(y_{t-1} - 4) + \varepsilon_t,$$

where $E(\varepsilon_t) = 0$. The most recent observation is

$$y_T = 8.$$

1. What are the unconditional mean μ and the AR coefficient ϕ? Is the process stationary?

2. Forecast inflation one and two periods ahead: $\widehat{y}_{T+1|T}$ and $\widehat{y}_{T+2|T}$.

3. What happens to the forecast as the forecast horizon becomes large? Explain briefly.

2 Midterm Review

Review Sheet

The following is a short summary of important formulas and concepts covered so far. It is intended for review and may not be comprehensive.

Do not bring this sheet to the exam. Enjoy! ♡

2.1 Basic Statistics and Probability

Sample mean:

$$\bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i.$$

Sample variance:

$$s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i - \bar{y})^2.$$

Sample standard deviation:

$$s = \sqrt{s^2}.$$

Cumulative distribution function (CDF):

$$F(y) = \mathrm{P}(Y \leq y).$$

Probability of an interval:

$$\mathrm{P}(a < Y \leq b) = F(b) - F(a).$$

Quantile function:

$$Q(p) = F^{-1}(p).$$

Central 90% interval:

$$[Q(0.05), Q(0.95)].$$

Central 95% interval:

$$[Q(0.025), Q(0.975)].$$

2.2 Normal Distribution

Normal distribution:

$$Y \sim N(\mu, \sigma^2).$$

Standardization:

$$Z = \frac{Y - \mu}{\sigma} \sim N(0, 1).$$

Normal CDF:

$$\mathrm{P}(Y \leq y) = \Phi\left(\frac{y-\mu}{\sigma}\right).$$

Normal quantile:

$Q_Y(p) = \mu + \sigma\Phi^{-1}(p)$.

Useful standard normal quantiles:

$\Phi^{-1}(0.95) \approx 1.645$.

$\Phi^{-1}(0.975) \approx 1.96$.

Central 90% interval:

$\mu \pm 1.645\sigma$.

Central 95% interval:

$\mu \pm 1.96\sigma$.

2.3 Forecast Evaluation

Forecast error:

$e = y - \widehat{y}$.

A positive forecast error means underprediction; a negative forecast error means overprediction.

Mean squared prediction error (MSPE):

$$\mathrm{MSPE} = \frac{1}{n}\sum_{i=1}^{n} e_i^2.$$

Root mean squared prediction error (RMSPE):

$\mathrm{RMSPE} = \sqrt{\mathrm{MSPE}}$.

Mean absolute prediction error (MAPE):

$$\mathrm{MAPE} = \frac{1}{n}\sum_{i=1}^{n} |e_i|.$$

Optimal forecast under squared-error loss:

$\widehat{y} = E(y \mid I)$.

Optimal forecast under absolute-error loss:

$\widehat{y} =$ conditional median of y given I.

2.4 Regression

Linear regression model:

$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i.$

Conditional mean assumption:

$E(\varepsilon_i \mid x_i) = 0.$

Conditional mean of y_i:

$E(y_i \mid x_i) = \beta_0 + \beta_1 x_i.$

Predicted value:

$\widehat{y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 x_i.$

Test statistic for $H_0 : \beta_j = b$:

$$t = \frac{\widehat{\beta}_j - b}{\mathrm{SE}(\widehat{\beta}_j)}.$$

Two-sided 5% test:

Reject H_0 if $|t| > 1.96$.

95% confidence interval for β_j:

$\widehat{\beta}_j \pm 1.96 \ \mathrm{SE}(\widehat{\beta}_j).$

2.5 Forecast Intervals

Normally distributed forecast error:

$e = y - \widehat{y} \sim N(0, v^2).$

90% forecast interval:

$\widehat{y} \pm 1.65v.$

95% forecast interval:

$\widehat{y} \pm 1.96v.$

Mean of past forecast errors:

$$\bar{e} = \frac{1}{n} \sum_{i=1}^{n} e_i.$$

Estimated standard deviation of forecast errors:

$$\widehat{v} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (e_i - \bar{e})^2}.$$

Higher coverage $\Rightarrow$ wider forecast interval.

2.6 Growth Rates and Real Values

Quarter-over-quarter (QoQ) growth:

$100(\ln Y_t - \ln Y_{t-1})$.

Annualized QoQ growth:

$400(\ln Y_t - \ln Y_{t-1})$.

Year-over-year (YoY) growth for quarterly data:

$100(\ln Y_t - \ln Y_{t-4})$.

Month-over-month (MoM) growth:

$100(\ln Y_t - \ln Y_{t-1})$.

Annualized MoM growth:

$1200(\ln Y_t - \ln Y_{t-1})$.

Year-over-year (YoY) growth for monthly data:

$100(\ln Y_t - \ln Y_{t-12})$.

Real value in index-base-year dollars:

$$\text{Real value} = \text{Nominal value} \times \frac{100}{\text{Price index}}.$$

2.7 Trend and Seasonality

Time series decomposition:

$y_t = T_t + S_t + C_t$.

Seasonal adjustment:

$y_t^{SA} = y_t - \widehat{S}_t$.

Linear trend:

$T_t = \beta_0 + \beta_1 t$.

Quadratic trend:

$T_t = \beta_0 + \beta_1 t + \beta_2 t^2$.

Detrended component:

$\widehat{C}_t = y_t - \widehat{T}_t$.

Forecast with trend and cycle:

$\widehat{y}_{t+h|t} = \widehat{T}_{t+h} + \widehat{C}_{t+h|t}$.

If C_t is i.i.d. with mean zero:

$\widehat{C}_{t+h|t} = 0$.

2.8 Seasonal Dummies

Quarterly seasonal model using four dummies and no intercept:

$y_t = \gamma_1 D_{1t} + \gamma_2 D_{2t} + \gamma_3 D_{3t} + \gamma_4 D_{4t} + C_t.$

Quarterly seasonal model using an intercept and three dummies:

$y_t = b_0 + b_2 D_{2t} + b_3 D_{3t} + b_4 D_{4t} + C_t.$

Intercept in the three-dummy representation:

$b_0 = \gamma_1.$

Seasonal dummy coefficients in the three-dummy representation:

$b_j = \gamma_j - \gamma_1, \qquad j = 2, 3, 4.$

Forecasted seasonal component if $t + h$ is in quarter j:

$\widehat{S}_{t+h} = \widehat{\gamma}_j.$

Forecast if $t + h$ is in quarter j:

$\widehat{y}_{t+h|t} = \widehat{\gamma}_j + \widehat{C}_{t+h|t}.$

2.9 AR(1) Model

AR(1) model centered around its mean:

$y_t - \mu = \phi(y_{t-1} - \mu) + \varepsilon_t.$

Equivalent AR(1) representation:

$y_t = m + \phi y_{t-1} + \varepsilon_t.$

Relationship between the intercept and the unconditional mean:

$m = (1 - \phi)\mu.$

Unconditional mean:

$\mu = \dfrac{m}{1 - \phi}.$

Stationarity condition:

$|\phi| < 1.$

One-step-ahead forecast:

$\widehat{y}_{T+1|T} = \mu + \phi(y_T - \mu).$

h-step-ahead forecast:

$\widehat{y}_{T+h|T} = \mu + \phi^h(y_T - \mu).$

Long-run forecast under stationarity:

$\lim_{h \to \infty} \widehat{y}_{T+h|T} = \mu.$

Answers

Exercise 1 Forecasting with an AR(1).

Suppose inflation follows

$$y_t - 4 = 0.5(y_{t-1} - 4) + \varepsilon_t,$$

where $E(\varepsilon_t) = 0$, and $y_T = 8$.

1. What are the unconditional mean μ and the AR coefficient ϕ? Is the process stationary?

   From the mean-centered AR(1),

   $$\mu = 4, \qquad \phi = 0.5.$$

   Since

   $$|\phi| = 0.5 < 1,$$

   the process is stationary.

2. Forecast inflation one and two periods ahead: $\widehat{y}_{T+1|T}$ and $\widehat{y}_{T+2|T}$.

   The one-step-ahead forecast is

   $$\widehat{y}_{T+1|T} = 4 + 0.5(8 - 4) = 6.$$

   Using the chain rule of forecasting,

   $$\widehat{y}_{T+2|T} = 4 + 0.5(6 - 4) = 5.$$

   Equivalently,

   $$\widehat{y}_{T+2|T} = 4 + (0.5)^2(8 - 4) = 5.$$

3. What happens to the forecast as the forecast horizon becomes large? Explain briefly.

   The h-step-ahead forecast is

   $$\widehat{y}_{T+h|T} = 4 + (0.5)^h(8 - 4).$$

   Since

   $$(0.5)^h \to 0,$$

   the forecast approaches

   $$\widehat{y}_{T+h|T} \to 4.$$

   Thus, forecasts revert toward the unconditional mean of 4 as the forecast horizon increases.