Forecasting

Consider two random variables Y & X with joint density function $f(x,y)$. Given the value of $X = x$, we wish to predict or forecast the value of Y using <u>any</u> function of X, say $h(X)$, as our predictor

Q: What is the <u>best</u> predictor; i.e. optimal function $h(\cdot)$?

A: First we must define what "best" predictor means.

A common measure of forecast accuracy is mean-square error

$$MSE(Y, h(X)) = E[(Y - h(X))^2]$$

= expected squared forecast error.

Now, we ask: "What $h(X)$ produces the smallest value of $MSE(Y, h(X))$?"

<u>Result:</u> $h(X) = E[Y \mid X=x]$ = conditional expectation of Y given X

Proof:

Let $h(X)$ be any function of X. Then

$MSE(Y, h(X)) = E[(Y-h(X))^2]$. Now

add & subtract $E[Y|X]$ inside the $(\ )^2$:

$$MSE(Y,h(X)) = E[((Y-E[Y|X])+(E[Y|X]-h(X)))^2]$$

$$= E[(Y-E[Y|X])^2 + 2(Y-E[Y|X])(E[Y|X]-h(X)) + (E[Y|X]-h(X))^2]$$

$$= E[(Y-E[Y|X])^2]$$

$$+ 2E[(Y-E[Y|X])(E[Y|X]-h(X))]$$

$$+ E[(E[Y|X]-h(X))^2]$$

Now use iterated expectations on middle piece, i.e.

$$E[\cancel{Y}] = E_X[E[Y|X]]$$

to give

$$E[(Y-E[Y|X])(E[Y|X]-h(X))] = E_X\{E[(Y-E[Y|X])(E[Y|X]-h(X))|$$

$$= E_X\left\{ \left(E\{Y|X\} - E\{Y|X\}\right)\left(E\{Y|X\} - h(X)\right) \right\}$$

$$= E_X\{0\} = 0$$

Hence,

$$MSE(Y, h(X)) = E\left[(Y - E\{Y|X\})^2\right] + E\left[(E\{Y|X\} - h(X))^2\right]$$

$$= MSE(Y, E\{Y|X\}) + \underbrace{E\left[(E\{Y|X\} - h(X))^2\right]}_{\text{always} \geq 0}$$

Clearly, $MSE(Y, h(X))$ is minimized using

$$h(X) = E\{Y|X\}$$

since, in this case, $E\left\{(E\{Y|X\} - h(X))^2\right\} = 0$.

Forecasting with the AR(1)

$$y_t = c + \phi y_{t-1} + \epsilon_t \qquad t = 1, \ldots, T$$

$$\epsilon_t \sim iid\ (0, \sigma^2)$$

$$|\phi| < 1 \quad : \quad E[Y_t] = \mu = \frac{c}{1-\phi}$$

$$var(Y_t) = \frac{\sigma^2}{1-\phi^2}$$

Goal: Given information at time T, $I_T = \{y_T, y_{T-1}, \ldots y_1\}$

and known values for c & ϕ, compute forecasts

of $y_{T+1}, y_{T+2}, \ldots$

Forecasting rule: $\hat{y}_{T+1} = E[y_{T+1} | I_T]$ = conditional expectation of y_{T+1} given information available at time T

$$\hat{y}_{T+2} = E[y_{T+2} | I_T]$$

$$\vdots$$

$$\hat{y}_{T+s} = E[y_{T+s} | I_T]$$

For the AR(1), these conditional expectations can be easily computed.

$s=1$: $\quad y_{T+1} = c + \phi y_T + \epsilon_{T+1}$

$$E[y_{T+1} | I_T] = c + \phi E[y_T | I_T] + E[\epsilon_{T+1} | I_T]$$

$$\Rightarrow \underbrace{E[y_{T+1} \mid I_T]}_{\hat{y}_{T+1}} = c + \phi y_T \qquad \text{since} \qquad E[\epsilon_{T+1} \mid I_T] =$$

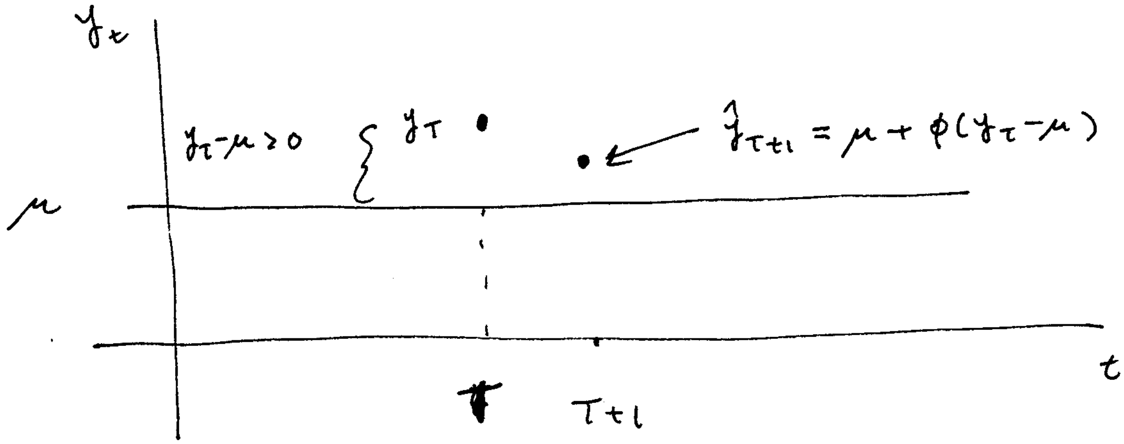
Alternatively, since $c = \mu(1-\phi)$ we have

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

$$= \mu + \phi(\underbrace{y_T - \mu}_{\substack{\text{deviation of} \\ y_T \text{ from} \\ \text{unconditional} \\ \text{mean.}}})$$

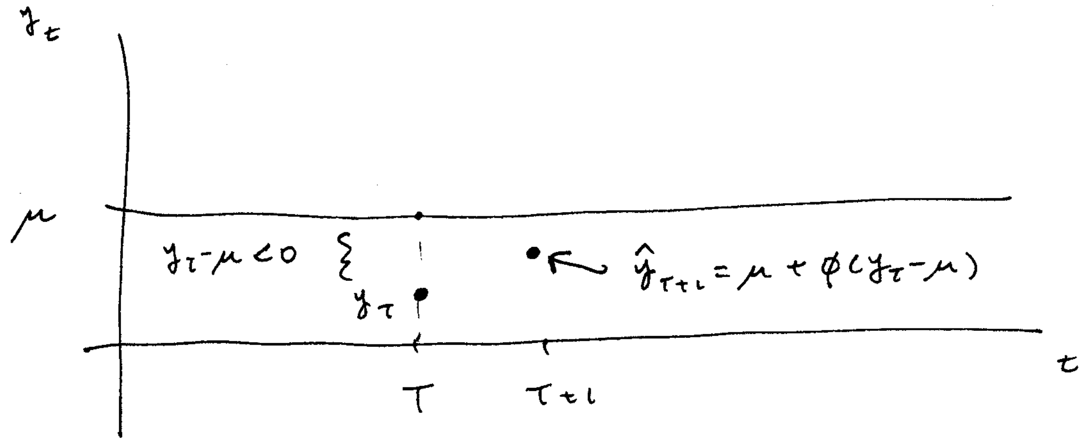
unconditional mean

If $\phi > 0$ and $y_T - \mu > 0$ then



Similarly, if $\phi > 0$ and $y_T - \mu < 0$ then

over ↓



We see that the forecasts are reverting to the mean of y_t.

~~See: $y_{T+2} = c + \phi y_{T+1} + \epsilon_{T+2}$~~

The error in the forecast is

$$\hat{\epsilon}_{T+1} = y_{T+1} - \hat{y}_{T+1}$$

$$= y_{T+1} - (c + \phi y_T)$$

$$= \epsilon_{T+1}$$

which has mean zero (Forecast is unbiased!) and variance

$$\mathrm{var}(\hat{\epsilon}_{T+1}) = \mathrm{var}(\epsilon_{T+1}) = \sigma^2.$$

A 95% confidence interval (approx) for y_{T+1}

Can then be computed as

$$\hat{y}_{T+1} \pm 2 * \sqrt{var(\hat{\epsilon}_{T+1})}$$

<u>Note:</u> $var(\hat{\epsilon}_{T+1}) = \sigma^2 < var(y_t) = \frac{\sigma^2}{1-\phi^2}$.

variance of forecast error using $\bar{y}$ as predictor.

$s=2$: $y_{T+2} = c + \phi y_{T+1} + \epsilon_{T+2}$

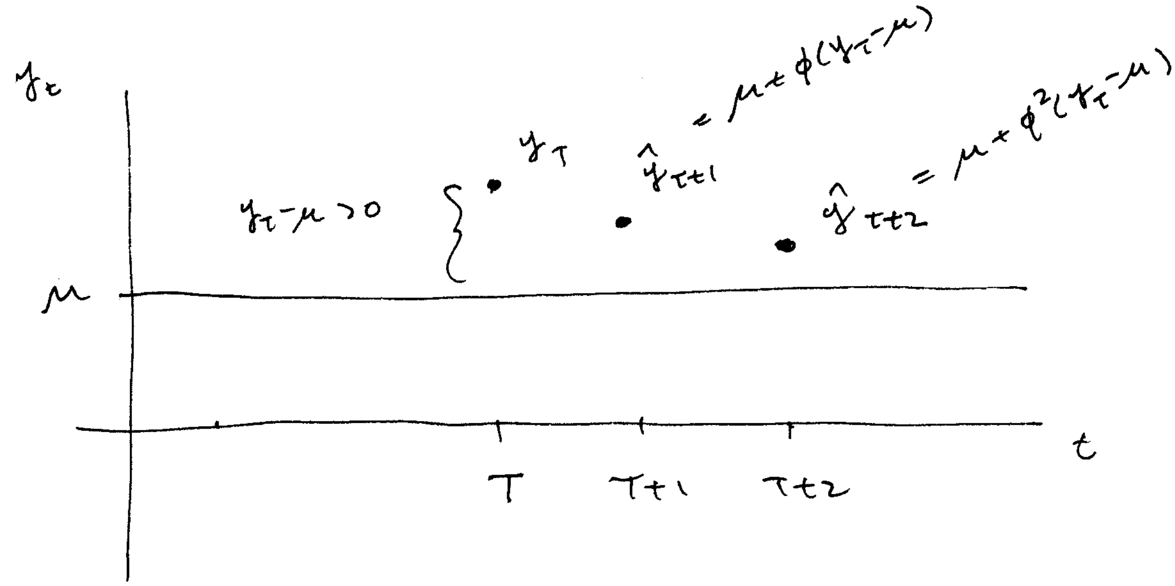
$$\hat{y}_{T+2} = E[y_{T+2} \mid I_T] = c + \phi E[y_{T+1} \mid I_T] + E[\epsilon_{T+2} \mid I_T]$$
$$= c + \phi \hat{y}_{T+1}$$

where $\hat{y}_{T+1} = \mu + \phi(y_T - \mu)$. Alternatively, we have

$$\hat{y}_{T+2} = \mu(1-\phi) + \phi[\mu + \phi(y_T - \mu)]$$
$$= \mu + \phi^2(y_T - \mu)$$

Graphically, if $\phi > 0$ ~~are true~~ and $(y_T - \mu) > 0$

we have



and we see that the forecast is reverting even close to the mean.

The error of the forecast is

$$\hat{\epsilon}_{T+2} = y_{T+2} - \hat{y}_{T+2}$$

$$= c + \phi y_{T+1} + \epsilon_{T+2} - [c + \phi \hat{y}_{T+1}]$$

$$= \phi [y_{T+1} - \hat{y}_{T+1}] + \epsilon_{T+2}$$

$$= \phi \hat{\epsilon}_{T+1} + \epsilon_{T+2}$$

$$= \phi \epsilon_{T+1} + \epsilon_{T+2} \quad \text{since} \quad \hat{\epsilon}_{T+1} = \epsilon_{T+1}$$

and

$$\text{var}(\hat{\epsilon}_{T+2}) = (\phi^2 + 1)\sigma^2$$

Notice that $\text{var}(\hat{\epsilon}_{T+2}) > \text{var}(\hat{\epsilon}_{T+1})$.

It should not be too hard to see that for general s

$$\hat{y}_{T+s} = E[y_{T+s} | I_T]$$

$$= c + \phi \, \hat{y}_{T+s-1}$$

$$= \mu + \phi^s (y_T - \mu)$$

<u>Note:</u> $\lim_{s \to \infty} \hat{y}_{T+s} = \mu$ provided $|\phi| < 1$

Also,

$$\hat{\epsilon}_{T+s} = y_{T+s} - \hat{y}_{T+s}$$

$$= \phi \, \hat{\epsilon}_{T+s-1} - \epsilon_{T+s}$$

$$= \phi^{s-1} \epsilon_{T+1} + \cdots + \phi \, \epsilon_{T+s-1} + \epsilon_{T+s}$$

and

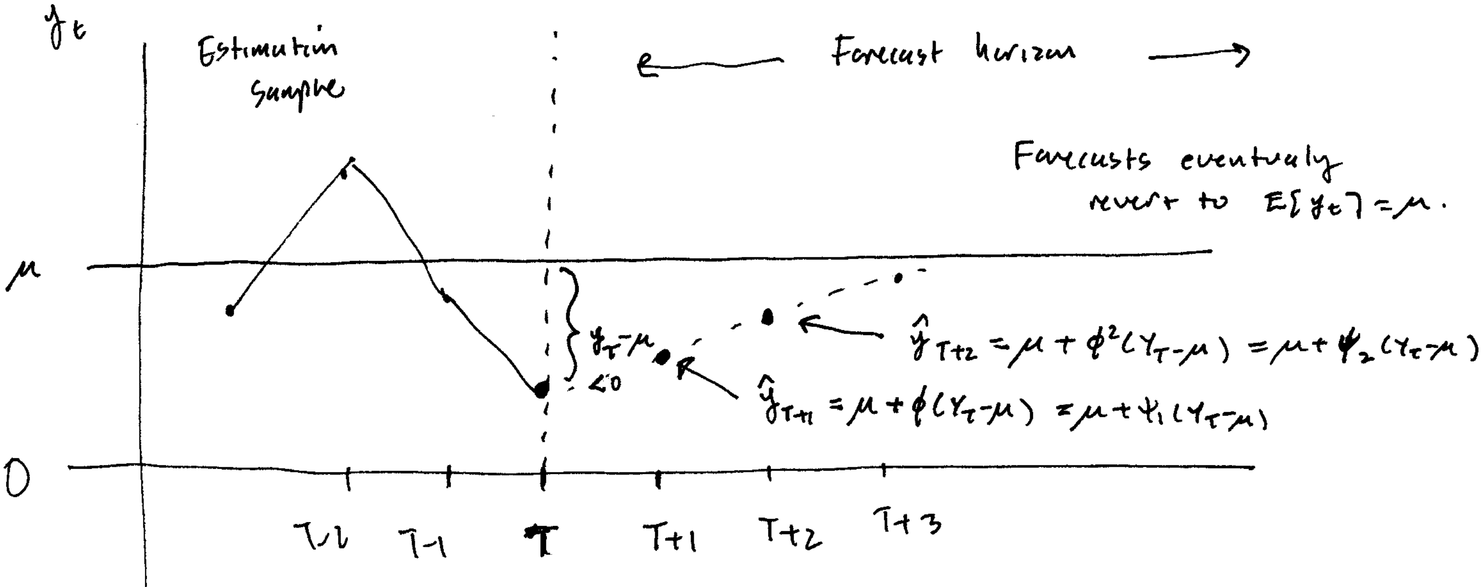
$$\mathrm{var}(\hat{\epsilon}_{T+s}) = \sigma^2 \left(1 + \phi^2 + \phi^4 + \cdots + (\phi^{s-1})^2 \right)$$

<u>Note:</u> $\lim_{s \to \infty} \mathrm{var}(\hat{\epsilon}_{T+s}) = \sigma^2 \sum_{k=0}^{\infty} \phi^{2k} = \dfrac{\sigma^2}{1-\phi^2}$

$$= \mathrm{var}(y_t) \; !$$

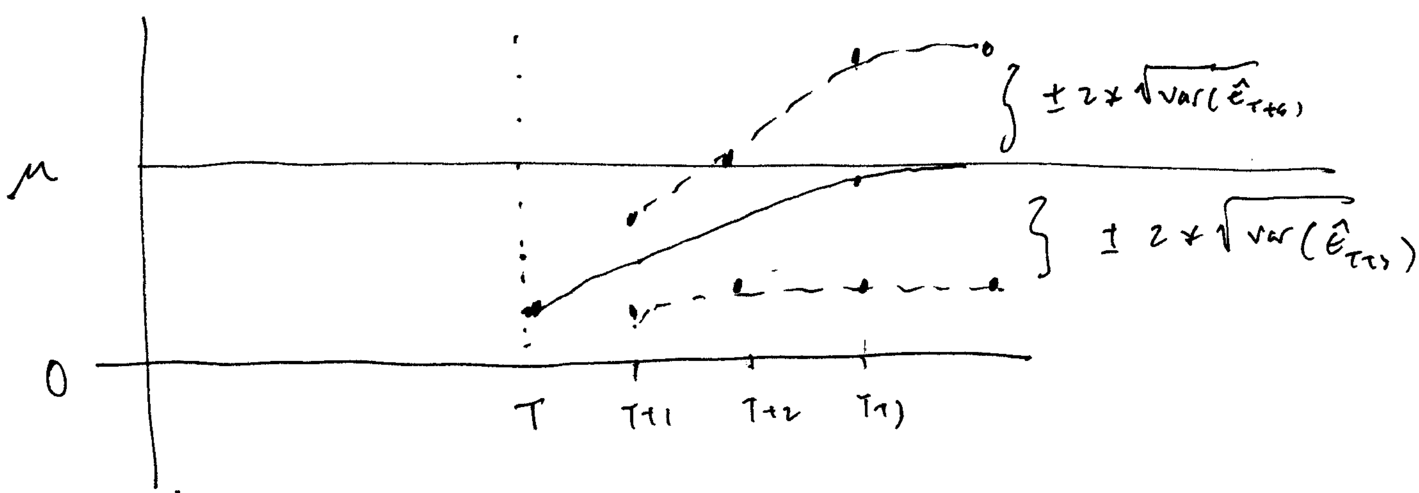
Graphically we have

$$y_t = c + \phi y_{t-1} + \epsilon_t \quad \text{or} \quad (y_t - \mu) = \phi(y_{t-1} - \mu) + \epsilon_t$$



<u>Note:</u> the path taken by the forecasts mimics the impulse response function of the process.

The forecasts usually are reported with forecast std. errors in the form of 95% Confidence intervals.



Example

Forecasting growth rate of real GDP using AR(1) model

$$y_t = \ln(RGDP_t) - \ln(RGDP_{t-1}) \approx \%\Delta RGDP_t.$$

Fitted AR(1) by OLS 1947:3 – 1992:4

$$\hat{y}_t = \underset{(0.0009)}{0.0048} + \underset{(0.0692)}{0.3737}\, y_{t-1}$$

$$\hat{\sigma} = 0.00945, \quad R^2 = 0.139$$

Unconditional Mean (quarterly) $= \dfrac{0.0048}{1 - 0.3737} = 0.0077$

>> Annual rate $\approx$ 3.1% per year.

Compute Forecasts of y_t over 1993 – 4 quarters.

$$\hat{y}_{1993:I} = 0.0077 + 0.3737\left[\, y_{1992:IV} - 0.0077 \right]$$

$$= 0.0077 + 0.3737\left[\underbrace{0.0139 - 0.0077}_{\text{above mean.}}\right]$$

$$= 0.0101$$

$$\hat{y}_{1993:II} = 0.0077 + 0.3737\left[\hat{y}_{1993:I} - 0.0077\right]$$

$$= 0.0077 + 0.3737\left[0.0101 - 0.0077\right]$$

$$= 0.0086$$

$$\hat{y}_{1993:III} = 0.0077 + 0.3737\left[\hat{y}_{1993:II} - 0.0077\right]$$

$$= 0.0077 + 0.3737\left[0.0086 - 0.0077\right]$$

$$= 0.0080$$

$$\hat{y}_{1993:IV} = 0.0077 + 0.3737\left[\hat{y}_{1994:III} - 0.0077\right]$$

$$= 0.0077 + 0.3737\left[0.0080 - 0.0077\right]$$

$$= 0.0078$$

So by 1993:IV the forecast of y_t is essentially the unconditional mean.

Forecast std. errors are usually computed assuming the estimated values of c & ϕ are the true values. Then

$$\hat{\text{var}}(\hat{\varepsilon}_{1993:I}) = \hat{\text{var}}(y_{1993:1} - \hat{y}_{1993:1}) = \hat{\sigma}^2 = (0.009457)^2$$

$$\Rightarrow SD(\hat{\varepsilon}_{1993:I}) = 0.009457$$

$y_{1993:I}$

A 95% CI. for ~~the forecast~~ is then

$$\hat{y}_{1993:1} \pm 2 * SD(\hat{\epsilon}_{1993:I})$$

$$= 0.0101 \pm 2*(0.009457)$$

$$= [-0.0088, 0.0290]$$

The actual value is $y_{1993:1} = 0.002881$ which is inside the 95% C.I.

For $\hat{y}_{1993:II}$, the estimated forecast error variance is

$$\hat{\text{var}}(\hat{\epsilon}_{1993:II}) = (\hat{\phi}^2+1)\cdot\hat{\sigma}^2$$

$$= ((0.3737)^2+1)*(0.009457)^2$$

$$= 0.0001019$$

$$\Rightarrow SD(\hat{\epsilon}_{1993:II}) = 0.0101 > SD(\hat{\epsilon}_{1993:I})$$

A 95% C.I for the future value is then

$$\hat{y}_{1993:II} \pm 2 * SD(\hat{\epsilon}_{1993:II}) = [-0.0116, 0.0288]$$

and the actual value $y_{1993:II} = 0.0059$ lies in the interval.