Autoregressive models, mean reversion and forecast accuracy

Quantitative Methods, Level Two

1 A story: yesterday helps predict tomorrow

- Sales growth carries some momentum
- Autoregressive: a series explained by its own past
- The past values used are called lags
- The number of lags is the order, p

IN THE LESSON

- Maya Ortega covers a made up supermarket chain called Tidewell Grocers.
- She wants to forecast its quarterly sales growth.
- She notices something.
- When growth was strong last quarter, it tends to be fairly strong this quarter too.
- Sales do not jump around at random.
- They carry momentum, and they also fade back.
- So Maya has an idea: use the series' own past to predict its future.
- A model that does this is called autoregressive.
- Auto means self, and regressive means we use regression.
- The series is regressed on its own earlier values.
- The number of past values used is called the order of the model, and we write it as p.

2 The autoregressive model of order p

$$x(t) = b_0 + b_1x(t-1) + b_2x(t-2) + \ldots + b_px(t-p) + \varepsilon$$

Value this period	The value we forecast
b zero: intercept	Starting level
b one to b p: lag weights	Weight on each lag
Past values: p lags	Earlier values of x
Error term	What the past misses

IN THE LESSON

- Here is the model.
- The value this period equals an intercept, plus a weight times the value one period ago, plus another weight times the value two periods ago, and so on, back p periods, plus an error.
- Take each piece.
- The intercept, b zero, is the starting level.
- Each b is the weight on one lag, and it is estimated by regression.
- The lags are the earlier values of the same series.
- The error term is the part of this period that the past cannot explain.
- If p is one, we call it an A R one model.
- If p is two, an A R two.
- A reminder from earlier: the series must be covariance stationary, or the estimates cannot be trusted.

3 Example 1: a one-lag model, three steps ahead

x(t) = 0.5 + 0.7·x(t−1). Latest growth: 4.0%

One step ahead	0.5 + 0.7 × 4.0	**3.3%**
Two steps ahead	0.5 + 0.7 × 3.3	**2.81%**
Three steps ahead	0.5 + 0.7 × 2.81	**2.47%**

IN THE LESSON

- Example one.
- Maya fits an A R one to Tidewell's sales growth, in per cent.
- The estimate is growth equals zero point five, plus zero point seven times last quarter's growth.
- The latest quarter came in at four per cent.
- For the one step forecast, we plug in the latest actual.
- Zero point five, plus zero point seven times four, is three point three.
- Now the next quarter.
- We do not yet know that quarter, so we use our forecast of three point three in its place.
- Zero point five, plus zero point seven times three point three, is two point eight one.
- And once more: zero point five, plus zero point seven times two point eight one, is about two point four seven.

4 The chain rule of forecasting

One step ahead: both lags are actual

First lag: latest, 3.0 → Second lag: before, 2.0 → Forecast: 2.7

Two steps ahead: first lag is the forecast

First lag: forecast 2.7 →
Second lag: latest, 3.0 → Forecast: 2.83

IN THE LESSON

- What we just did has a name: the chain rule of forecasting.
- Here it is drawn for a model with two lags.
- In the top row, we forecast one quarter ahead.
- Both lags are things that have already happened, so we use the actual values: the latest quarter and the one before.
- In the bottom row, we forecast two quarters ahead.
- Now the first lag is a quarter that has not happened yet, so the forecast we just made takes its place.
- The second lag is the latest actual quarter, which we know, so we keep it as it is.
- The old, earlier actual is no longer needed, because it is now too far back.
- Each step uses the one before, like links in a chain.
- The figures on the boxes come from the next example.

5 Example 2: a two-lag model, one and two steps

x(t) = 1.0 + 0.40·x(t−1) + 0.25·x(t−2). Latest 3.0%, before 2.0%

One step ahead	1.0 + 0.40×3.0 + 0.25×2.0	**2.70%**
Two steps ahead	1.0 + 0.40×2.7 + 0.25×3.0	**2.83%**
Slip: old value as second lag	1.0 + 0.40×2.7 + 0.25×2.0	**2.58%**

IN THE LESSON

- Example two, with two lags.
- Maya's second model says growth equals one, plus zero point four times the last quarter, plus zero point two five times the quarter before that.
- The latest quarter was three per cent and the one before it was two per cent.
- One step ahead, both lags are actual.
- One, plus zero point four times three, plus zero point two five times two.
- That is one, plus one point two, plus zero point five, which is two point seven per cent.
- Two steps ahead, the first lag is our forecast, two point seven, and the second lag is the latest actual, three.
- One, plus zero point four times two point seven, plus zero point two five times three.
- That is one, plus one point oh eight, plus zero point seven five, which is two point eight three per cent.

6 Reading the forecast: less certain at every step

- Each step ahead, the lags move forward too
- Add the intercept again at every step
- Use actual values wherever they are known
- Further ahead means less certain

IN THE LESSON

- Notice two things about the answer.
- First, the slip on the last row.
- A common error is to keep the old two per cent as the second lag in the two step forecast.
- But once we move one quarter forward, the lags move with us.
- The old value is now three quarters back, outside this model.
- Second, notice the intercept appears in every step.
- It is part of every use of the equation.
- The last point is about trust.
- The one step forecast is uncertain because of the error term and the estimated weights.
- The two step forecast has all of that, and it also rests on a number that was itself a forecast.
- So uncertainty grows with every step ahead.

7 A story: reading the leftovers

- Residuals: what the model could not explain
- Good model: residuals look like random noise
- Autocorrelated residuals: a pattern was missed
- So we test them, lag by lag

IN THE LESSON

- Now Maya has a model.
- How does she know it is any good?
- Think of a tailor who makes a suit.
- If the cloth was cut well, the leftover scraps are just random offcuts.
- If the scraps are all long thin strips, the tailor missed a pattern.
- A regression has leftovers too: the residuals, which are what the model could not explain.
- If the model has captured the pattern in the series, the residuals should be random noise.
- In particular, today's residual should not be predictable from yesterday's.
- When residuals are correlated with their own past values, that is called residual autocorrelation, and it means the model is missing something it should have used.

8 Why not the Durbin Watson test?

- A lagged dependent variable is a regressor
- That makes Durbin Watson invalid here
- Test the residual autocorrelations instead
- One test per lag: is it different from zero?

IN THE LESSON

- You may remember the Durbin Watson statistic from regression, which tests for serial correlation in the residuals.
- It is tempting to use it here.
- But in an autoregressive model, one of the regressors is a lagged value of the dependent variable.
- That breaks the assumptions behind the Durbin Watson test, so its result is not valid.
- Instead the curriculum has us look at the autocorrelations of the residuals directly.
- We compute the correlation between the residuals and their own values one period back, two periods back, three, and so on.
- Then we ask, for each lag, whether that autocorrelation is statistically different from zero.
- That needs a standard error, which is the next slide.

9 Testing a residual autocorrelation

$t = \rho \div SE$ where $SE = 1 \div \sqrt{T}$, so $t = \rho \times \sqrt{T}$

Residual autocorrelation	Correlation with a lag
T: observations	Number of observations
Standard error	One over root T
t-statistic	Compare with about 2

IN THE LESSON

- Here is the test.
- If the model is right, each residual autocorrelation has a standard error of one divided by the square root of T, where T is the number of observations in the series.
- The t-statistic is the autocorrelation divided by that standard error.
- Dividing by one over root T is the same as multiplying by root T, so the t-statistic is simply the autocorrelation times the square root of T.
- We compare it with the critical value for the test, which for a decent sized sample at the five per cent level is about two.
- If the t-statistic is bigger than that in size, the autocorrelation is significant, and the model is missing information.

10 Example 3: Maya's residual check

AR(1) on 64 quarters: residual autocorrelations and $t = \rho \times 8$

Lag and residual autocorrelation	t-statistic
Lag 1: 0.06	0.48
Lag 2: −0.10	−0.80
Lag 3: 0.09	0.72
Lag 4: 0.31	2.48

IN THE LESSON

- Example three.
- Maya fits an A R one to sixty four quarters of Tidewell's sales growth and collects the residuals.
- Here are the residual autocorrelations for lags one to four.
- At lag one, the autocorrelation is zero point oh six.
- At lag two, minus zero point one oh.
- At lag three, zero point oh nine.
- At lag four, zero point three one.
- Each is only a number.
- To tell which matter, we turn each one into a t-statistic, shown beside it.
- With sixty four observations the standard error is one over eight, so each t-statistic is the autocorrelation times eight.
- Next we see them as bars, and then we check the arithmetic.

11 Only one lag stands out

Absolute t-statistic

Lag 1	**0.48**
Lag 2	**0.80**
Lag 3	**0.72**
Lag 4	**2.48**

Critical value is about 2.0

IN THE LESSON

- Here are those t-statistics as bars, ignoring the sign, since a large negative value matters just as much as a large positive one.
- The critical value is about two.
- The first three bars are all well below it: zero point four eight, zero point eight oh and zero point seven two.
- Those lags are fine.
- The fourth bar is different.
- It reaches two point four eight, which is beyond two.
- So the lag four residual autocorrelation is significant.
- In quarterly data, lag four means the same quarter a year earlier.
- Something seasonal is left in the residuals.
- Notice that if Maya had only checked lag one, the model would have looked perfectly healthy.
- The check must cover several lags.

12 Example 3: the numbers behind the bars

T = 64 quarters. Lag 4 autocorrelation = 0.31

Standard error	1 ÷ √64 = 1 ÷ 8	**0.125**
Lag 4 t-statistic	0.31 ÷ 0.125	**2.48**
Critical value (as given)	about 2.0 at 5%	**2.0**
Smallest significant size	2.0 × 0.125	**0.25**
Verdict	lag 4 is significant	**add a lag**

IN THE LESSON

- Let us do the arithmetic once, step by step.
- The standard error is one over the square root of sixty four.
- The square root of sixty four is eight, so the standard error is zero point one two five.
- The lag four t-statistic is zero point three one divided by zero point one two five, which is two point four eight.
- Our critical value, as given in this example, is about two.
- Two point four eight is larger, so the autocorrelation is significant.
- Here is a useful shortcut.
- With this standard error, an autocorrelation needs to be bigger than about two times zero point one two five, or zero point two five, to be significant.
- Lag four, at zero point three one, clears that bar.

13 What to do when a lag is significant

Repeat until the residuals are clean

Fit the model → Test residual autocorrelations → Add a lag if significant → Refit and test again

IN THE LESSON

- So what does Maya do?
- A significant residual autocorrelation means the model is missing information, so she changes the model and tests again.
- The usual repair is to add a lag.
- If the significant lag is four in quarterly data, or twelve in monthly data, it points to seasonality, so she adds that seasonal lag.
- She adds lag four to her model, refits it, and checks the new residuals.
- In our example, the new residual autocorrelations are all small, and none has a t-statistic above about one.
- Then the model passes.
- A last warning: passing this check does not finish the job.
- We still need to test the series for stationarity, and the residuals for conditional heteroskedasticity.

14 A story: the thermostat

- Mean reverting: falls when above, rises when below
- Like a room pulled back to the thermostat setting
- The target is the mean reverting level
- An AR(1) model gives it in one line

IN THE LESSON

- Think about a thermostat set to eighteen degrees.
- If the room is warm, the heating switches off and the room cools.
- If it is cold, the heating comes on and the room warms.
- The room is always being pulled back towards eighteen.
- Many financial series behave like this.
- A company's profit margin drifts away from its normal level, then competition, costs and demand pull it back.
- When a series tends to fall when it is above its mean, and rise when it is below its mean, we call it mean reverting.
- The mean it heads towards is the mean reverting level.
- For an A R one model, there is a one line formula for it.

15 The mean reverting level

Mean-reverting level = $b_0 \div (1 - b_1)$ (needs $b_1 < 1$)

b zero: intercept	Intercept
b one: slope on the lag	Slope on the lag
Mean-reverting level	Where forecasts head

IN THE LESSON

- Here is the formula.
- The mean reverting level equals b zero divided by one minus b one.
- Where does it come from?
- The level is the value at which the model predicts no change.
- So we set the value now equal to the value before, call it x, and write x equals b zero plus b one times x.
- Move the b one term across, and x times one minus b one equals b zero.
- Divide through, and you have the formula.
- It only works when b one is less than one in absolute size.
- The intercept is b zero, the slope on the lag is b one, and the answer is the level the series is pulled towards.
- Please remember it is a long run anchor, not the next forecast.

16 Example 4: Tidewell's gross margin

x(t) = 3.6 + 0.8·x(t−1). Gross margin, %

Mean-reverting level	3.6 ÷ (1 − 0.8)	**18%**
Check at 18	3.6 + 0.8 × 18	**18%**
From 21: next quarter	3.6 + 0.8 × 21	**20.4%**
From 21: the quarter after	3.6 + 0.8 × 20.4	**19.92%**
From 15: next quarter	3.6 + 0.8 × 15	**15.6%**

IN THE LESSON

- Example four.
- Maya models Tidewell's gross margin, in per cent, with an A R one: margin equals three point six, plus zero point eight times last quarter's margin.
- The mean reverting level is three point six divided by one minus zero point eight.
- That is three point six divided by zero point two, which is eighteen per cent.
- Check it: three point six, plus zero point eight times eighteen, is three point six plus fourteen point four, which is eighteen again.
- No change, as it should be.
- Now the margin is twenty one per cent, above the level.
- Next quarter's forecast is three point six plus zero point eight times twenty one, or twenty point four.
- Then nineteen point nine two.
- Falling towards eighteen.
- If it were fifteen, the forecast would be fifteen point six, rising.

17 Pulled towards eighteen from both sides

Chart: Gross margin (%) vs Quarters ahead

- 21%
- 15%
- Level 18%
- Falls from above
- Rises from below

IN THE LESSON

- Let us draw it.
- Time runs along the bottom, in quarters ahead, and the margin is up the side.
- The flat line is the mean reverting level, eighteen.
- Now follow the upper path.
- It starts at twenty one per cent and falls, quickly at first and then more slowly, getting closer to eighteen without quite touching it.
- The lower path starts at fifteen and rises the same way.
- Why does it curve?
- Each quarter the series closes the same share of the remaining gap, one minus b one, which here is twenty per cent.
- The gap from twenty one is three, then two point four, then one point nine two.
- As the gap shrinks, so does each step.

18 The random walk has no level

- $b_1 = 1$: the formula divides by zero
- A random walk has no mean-reverting level
- Stationary series always have a finite level
- Closer b_1 is to 1, the weaker the pull

IN THE LESSON

- What if b one is exactly one?
- Look at the formula.
- One minus one is zero, and dividing by zero has no answer.
- A series like this is a random walk.
- Today's value is yesterday's value plus a random shock, so there is no level it is pulled back to.
- It can wander off anywhere, and its forecast is simply its latest value.
- The link is worth knowing.
- Every covariance stationary series has a finite mean reverting level.
- A series without one, like a random walk, cannot be covariance stationary, so an ordinary autoregressive model would be a poor choice for it.
- As b one creeps closer to one, the level moves further away and the pull gets weaker.

19 Mid-lesson check: five slips to avoid

- Two-step forecast: use the step-one forecast as lag 1
- Add the intercept at every step
- No Durbin–Watson: test residual autocorrelations
- Level is $b_0 \div (1 - b_1)$, not $b_0 \div b_1$ or $\div (1 + b_1)$
- The level is a long-run anchor, not next period's forecast

IN THE LESSON

- Before we move on, here are five slips that cost marks.
- One.
- In a two step forecast, keeping the old value as a lag.
- Once you step forward, the one step forecast is the first lag, and the latest actual becomes the second.
- Two.
- Forgetting to add the intercept again at each step.
- Three.
- Using Durbin Watson to check an autoregressive model, when you should test the residual autocorrelations at several lags, with a standard error of one over root T.
- Four.
- Getting the mean reverting level wrong, by dividing b zero by b one, or by one plus b one.
- The right denominator is one minus b one.
- Five.
- Treating the level as next period's forecast.
- It is where the series is heading, not where it will be next.

20 A story: the student who memorised the answers

- In-sample error: how well the model fits its own data
- Out-of-sample error: how well it predicts unseen data
- Fitting history is easy; forecasting is the test
- More lags and variables fit better but can forecast worse

IN THE LESSON

- Imagine a student who memorises last year's exam paper, answer by answer.
- Asked about that same paper, the student scores full marks.
- Asked a fresh question, the student is lost.
- The model has learned the past, not the subject.
- Forecasting models can do exactly this.
- A forecast made within the sample is one for a period the model was estimated on, and the error is an in sample forecast error.
- It tells you how well the model fits its own data.
- A forecast made outside the sample is for a period the model never saw, and its error is an out of sample error.
- That is the true test, because real forecasting is always about data we have not yet seen.

21 Fit improves, forecasts worsen

RMSE (percentage points)

A in-sample	**0.80**
B in-sample	**1.00**
A out-of-sample	**1.58**
B out-of-sample	**1.02**

Model A: many lags. Model B: AR(1)

IN THE LESSON

- Maya now has two candidate models for Tidewell's sales growth.
- Model A is rich, with many lags and extra variables.
- Model B is a simple A R one.
- She estimates both on data up to the end of last year, then forecasts the four quarters of this year.
- Here are the root mean squared errors.
- Look at the first two bars, the in sample figures.
- Model A fits better: zero point eight, against one point oh for Model B.
- Now look at the last two, the out of sample figures.
- The picture reverses.
- Model A has one point five eight, while Model B has one point oh two.
- The richer model bent itself around the quirks of the past, and those quirks did not come back.

22 Root mean squared error

RMSE = √[Σ(error²) ÷ n]

Forecast error	Actual minus forecast
Squared error	Error times itself
n: number of forecasts	Count of forecasts, not n − 1
RMSE	Lower is better

IN THE LESSON

- So how is the root mean squared error built?
- The name tells you the recipe, read backwards.
- Take each forecast error, which is the actual value minus the forecast.
- Square it.
- Take the mean of the squares, which means add them up and divide by the number of forecasts, n.
- Then take the root.
- Why square?
- Errors above and below zero would otherwise cancel.
- Squaring makes every error positive, and it also gives large errors more weight than small ones, so one big miss hurts.
- The model with the lower out of sample root mean squared error is preferred.
- It is measured in the same units as the data, here percentage points of growth.

23 Example 5: Model A, out of sample

Model A forecast errors (actual − forecast), 4 quarters

Errors	+2.0, −2.0, +1.0, −1.0	**mean 0**
Squared errors	4, 4, 1, 1	
Sum of squares	4 + 4 + 1 + 1	**10**
Mean of the squares	10 ÷ 4	**2.5**
RMSE	√2.5	**1.58**

IN THE LESSON

- Example five.
- Here are Model A's errors for the four quarters it forecast: plus two, minus two, plus one and minus one percentage points.
- Step one, square each.
- Four, four, one and one.
- Step two, add them: ten.
- Step three, divide by the number of forecasts, four.
- Ten divided by four is two point five.
- Step four, take the square root.
- The square root of two point five is about one point five eight.
- Notice something about Model A.
- Its errors add to zero, so its average error is exactly zero.
- A table showing only the average error would make Model A look perfect.
- The errors are large, they just cancel out.

24 Example 5: Model B and the verdict

Model B forecast errors (actual − forecast), 4 quarters

Errors	+0.9, +1.1, +0.7, +1.3	**mean 1.0**
Squared errors	0.81, 1.21, 0.49, 1.69	
Mean of the squares	4.2 ÷ 4	**1.05**
RMSE	√1.05	**1.02**
Compare: A 1.58, B 1.02	lower out-of-sample RMSE wins	**Model B**

IN THE LESSON

- Now Model B.
- Its errors were plus zero point nine, plus one point one, plus zero point seven and plus one point three.
- Square each: zero point eight one, one point two one, zero point four nine and one point six nine.
- They add to four point two.
- Divide by four: one point oh five.
- Take the root, and the root mean squared error is about one point oh two.
- So which is preferred?
- Model B, with one point oh two against one point five eight, because lower is better.
- Yet Model B's errors are all positive, with an average of one point oh.
- If Maya looked only at the average error, she would choose the wrong model.
- Squaring is what exposes the large offsetting misses in Model A.

25 Traps when scoring forecasts

- Rank on out-of-sample RMSE, not in-sample
- Divide by n, never n − 1
- A small mean error can hide big offsetting errors
- RMSE punishes big misses more than mean absolute error

IN THE LESSON

- A few traps to avoid.
- First, ranking models on in sample error.
- As we saw, the richer model almost always wins that contest, and it tells you little about forecasting.
- Rank on out of sample error.
- Second, dividing by n minus one.
- That belongs to sample variance.
- Here you divide by n, the number of forecasts.
- For Model B, dividing by three would give about one point one eight, which is wrong.
- Third, trusting a small average error.
- Model A's was zero, hiding errors of up to two points.
- Fourth, confusing root mean squared error with the mean absolute error.
- The absolute version just averages the sizes of the errors, so it treats a big miss and a small miss the same way per point.

26 A story: the road that changed its speed limit

- Coefficients describe the period they were estimated on
- Policy, economy or business can change
- A new regime can make the old model forecast badly
- So test whether the coefficients are stable

IN THE LESSON

- Suppose you learned how long your commute takes by timing it for ten years.
- Then the council cuts the speed limit and adds traffic lights.
- Your old rule of thumb stops working, even though it was carefully measured.
- Time series models have the same problem.
- Their coefficients are estimated from a stretch of history, and they describe how the world worked then.
- When the economy, government policy or a company's business changes, the coefficients can change too.
- A model estimated on one period may then forecast badly in another.
- A change like that is called a regime change.
- This last outcome is about how to spot it, and what to do about it.

27 Testing for instability

Stability check for a model

Split the sample → Fit each subperiod → Compare coefficients → Significant gap: new regime

IN THE LESSON

- How do we test for instability?
- The idea is simple.
- Split the sample into subperiods, fit the same model to each, and compare the coefficients.
- If they are significantly different, that is evidence of a regime change.
- Maya tries it on a made up economy's inflation, using an A R one for two decades and then for both together.
- She splits the twenty years into an earlier decade and a recent decade.
- She fits the model to each.
- She then tests whether the slope and intercept differ.
- In our example, the test says the difference is significant.
- We are told that, rather than calculating it.
- Then comes the hard part: deciding which period best represents the future.

28 Example 6: one series, three fits

AR(1) fits to annual inflation (%). Level = $b_0 \div (1 - b_1)$

Fitted model	Reverts to
Earlier decade: 1.8 + 0.40x	3.0%
Recent decade: 0.6 + 0.70x	2.0%
Pooled, 20 years: 1.0 + 0.60x	2.5%

IN THE LESSON

- Here are Maya's three fits to annual inflation, in per cent.
- Each is an A R one, so each has a mean reverting level of b zero divided by one minus b one.
- Earlier decade: one point eight, plus zero point four times last year's inflation.
- Its level is one point eight divided by zero point six, which is three per cent.
- Recent decade: zero point six, plus zero point seven times last year.
- Its level is zero point six divided by zero point three, which is two per cent.
- Pooled over all twenty years: one, plus zero point six times last year.
- The level is one divided by zero point four, which is two and a half per cent.
- Notice the pooled level is a blend of the other two.

29 Three models, three destinations

Mean-reverting level (%)

Earlier decade	**3.0%**
Recent decade	**2.0%**
Pooled	**2.5%**

Pooling describes neither decade

IN THE LESSON

- Here are those three levels as bars, and the picture makes the problem clear.
- Inflation settles near three per cent in the earlier decade, and near two per cent in the recent decade.
- The pooled model says two and a half per cent.
- But inflation did not settle at two and a half in either decade.
- By pooling, we blended two different processes into one that describes neither of them well.
- That is the trap.
- The pooled model has the most data, so its standard errors will be the smallest, and it looks the most reliable.
- But precise answers to the wrong question are no use.
- More data helps only if the process behind it stayed the same.

30 Example 6: the forecasts disagree

Inflation this year: 2.6%. Forecast next year

Earlier decade	1.8 + 0.40 × 2.6	**2.84%**
Recent decade	0.6 + 0.70 × 2.6	**2.42%**
Pooled	1.0 + 0.60 × 2.6	**2.56%**
Spread of the forecasts	2.84 − 2.42	**0.42**

IN THE LESSON

- Now forecast.
- Inflation this year is two point six per cent.
- The earlier decade model gives one point eight, plus zero point four times two point six.
- That is two point eight four.
- The recent decade model gives zero point six, plus zero point seven times two point six, which is two point four two.
- The pooled model gives one, plus zero point six times two point six, which is two point five six.
- So the forecasts range from two point four two to two point eight four, a spread of zero point four two.
- Which should Maya use?
- There is no formula for this.
- She must judge which period best represents the future.
- If the change came from something lasting, such as a new policy framework, the recent model is the better guide.

31 The trade-off and the traps

- Long sample: precise, if the process is stable
- If it changed, pooling blends regimes
- Check stationarity in each subperiod
- Significant over the full sample does not prove stable

IN THE LESSON

- This creates a trade-off.
- A longer sample gives more precise estimates, if the process was stable.
- But if the process changed, pooling blends different regimes and describes none of them well.
- A shorter, recent sample is more relevant but noisier.
- The skill is judging which period best represents the future.
- There are two more points.
- First, check that the series is covariance stationary within each subperiod, before you compare coefficients.
- Second, do not be reassured by coefficients that are significant over the long sample.
- A coefficient can be highly significant overall and still describe neither half.
- Significance tells you the coefficient is not zero.
- It does not tell you it stayed the same.

32 Recap: five jobs for a forecasting model

Build it, check it, anchor it, score it, doubt it

Build and forecast	Chain rule, intercept each step
Check the residuals	t = ρ × √T, many lags
Mean reversion	$b_0 \div (1 - b_1)$
Judge accuracy	Lowest out-of-sample RMSE
Watch for regimes	Split the sample, compare

IN THE LESSON

- Let us put it together.
- Build and forecast: an autoregressive model uses its own lags, and the chain rule replaces unknown lags with forecasts, adding the intercept each time.
- Check the residuals: the autocorrelation at each lag has standard error one over root T, so t is the autocorrelation times root T, and a significant lag means add a lag or a seasonal lag.
- Mean reversion: b zero over one minus b one, the long run anchor, which does not exist for a random walk.
- Judge accuracy: the lowest out of sample root mean squared error wins, and you divide by n.
- And watch for regimes: coefficients drift, so compare subperiods and choose the most relevant period.

33 Try three quick questions

1. AR(1): x = 1.2 + 0.5·x(t−1), latest 4. Forecast and level?
   Answer: Forecast 3.2; level 1.2 ÷ 0.5 = 2.4.
2. T = 100, residual autocorrelation 0.15. Significant?
   Answer: t = 0.15 × 10 = 1.5, below 2: not significant.
3. Model 1 better in-sample, Model 2 better out-of-sample. Prefer?
   Answer: Model 2: out-of-sample RMSE decides.

IN THE LESSON

- Now try three quick questions yourself.
- Pause after each one.
- Question one.
- An A R one model is growth equals one point two, plus zero point five times last period, and the latest value is four.
- What is next period's forecast, and the mean reverting level?
- One point two plus zero point five times four is three point two.
- The level is one point two divided by zero point five, which is two point four.
- Question two.
- With one hundred observations, a residual autocorrelation is zero point one five.
- Is it significant, with a critical value of about two?
- The t-statistic is zero point one five times ten, which is one point five.
- That is not significant.
- Question three.
- Model one has the better in sample error, and Model two has the better out of sample error.
- Which do you prefer?
- Model two, because out of sample error decides.