Linear univariate modelling based on multivariate information

Laurent Ferrara

Master 2 EIMPC
Université Paris Ouest

Introduction

- Additional information can be useful to autoprojective processes
- Especially for macroeconomic forecasting
- What type of information?
  Coincident macro indicators: hard data (IPI, retail sales, consumption ...)
  Leading indicators : opinion surveys, financial variables, composite indicators (OECD, US CLI ...)

Correlations

How to measure the relationship between Y_t and X_t?

Basic tool: contemporeanous correlation

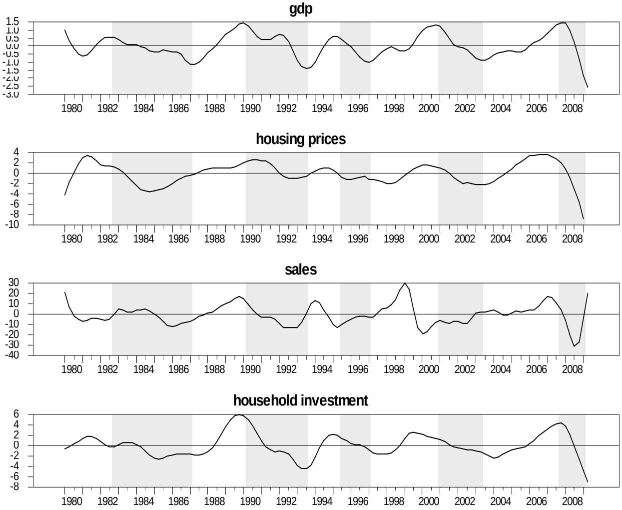
$$\rho(X_t, Y_t) = \frac{cov(X_t, Y_t)}{\sigma_X \sigma_Y}$$

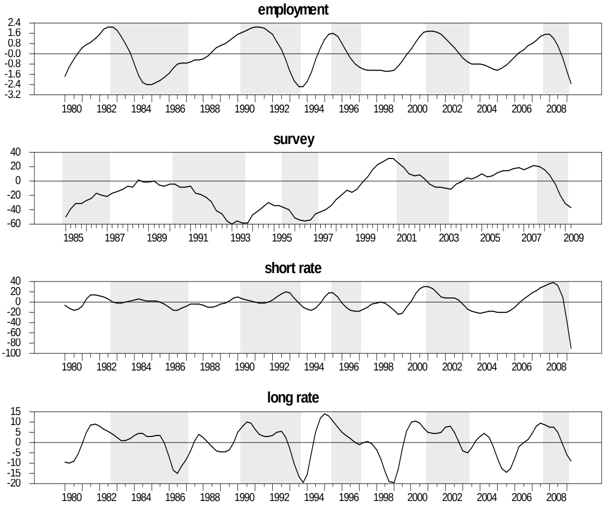
Alternative: cross-correlation at a given lag k

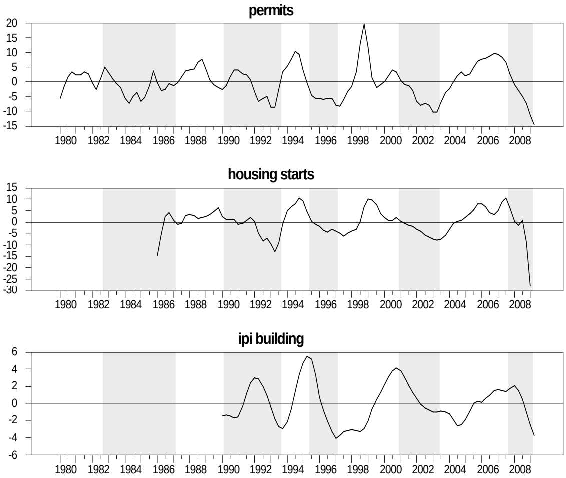
$$\rho(X_t, Y_{t-k}) = \frac{cov(X_t, Y_{t-k})}{\sigma_X \sigma_Y}$$

Example of cross-correlation: French housing market

- Objective: Etablish cyclical relationships between a set of macro and housing variables using correlation analysis
- Selected variables 1980q1 - 2009q2:
- Macro: GDP, Household investment, Employment in construction, IPI in construction
- Housing: Real prices, Sales, Permits, Starts, Survey by Property Developers
- Finance: Long (Gov. bonds 10 years) and Short (Euribor 3-months) interest rates


gdp
1.5
1.0
0.5
0.0
-0.5
-1.0
-1.5
-2.0
-2.5
-3.0
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
housing prices
4
2
0
-2
-4
-6
-8
-10
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
sales
30
20
10
0
-10
-20
-30
-40
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
household investment
6
4
2
0
-2
-4
-6
-8
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008


employment
2.4
1.6
0.8
-0.0
-0.8
-1.6
-2.4
-3.2
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
survey
40
20
0
-20
-40
-60
1985 1987 1989 1991 1993 1995 1997 1999 2001 2003 2005 2007 2009
short rate
40
20
0
-20
-40
-60
-80
-100
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
long rate
15
10
5
0
-5
-10
-15
-20
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008


permits
20
15
10
5
0
-5
-10
-15
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
housing starts
15
10
5
0
-5
-10
-15
-20
-25
-30
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008
ipi building
6
4
2
0
-2
-4
-6
1980 1982 1984 1986 1988 1990 1992 1994 1996 1998 2000 2002 2004 2006 2008

Correlation Analysis

	GDP	Prices	Sales	Invest.	Employ.	Survey	Short	Long	Permits	Starts	IPI
GDP	1	0.44	0.01	0.80	0.72	0.57	0.68	0.53	0.32	0.45	0.60
Prices	0.72	1	0.20	0.60	0.54	0.48	0.18	0.19	0.61	0.62	0.26
Sales	0.64	0.56	1	0.23	0.21	0.10	-0.37	-0.42	0.38	0.40	-0.40
Invest.	0.75	0.81	0.53	1	0.65	0.53	0.46	0.39	0.40	0.57	0.39
Employ.	0.79	0.75	0.54	0.79	1	0.46	0.56	0.64	0.17	0.22	0.57
Survey	0.80	0.80	0.67	0.81	0.71	1	0.28	0.21	0.39	0.40	0.31
Short	0.77	0.58	0.54	0.65	0.66	0.66	1	0.57	0.10	0.09	0.60
Long	0.74	0.73	0.44	0.69	0.69	0.59	0.66	1	-0.13	0.05	0.54
Permits	0.70	0.72	0.66	0.70	0.63	0.82	0.54	0.59	1	0.67	0.13
Starts	0.80	0.82	0.69	0.87	0.71	0.84	0.64	0.67	0.86	1	0.31
IPI	0.81	0.64	0.40	0.72	0.85	0.72	0.74	0.65	0.56	0.63	1

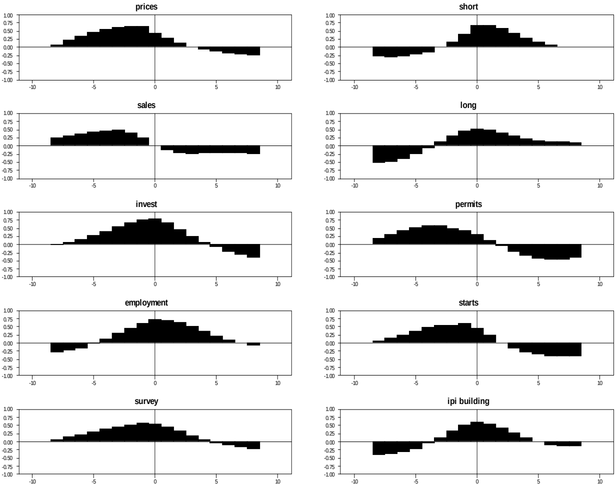
Table: Concordance indexes for contemporaneous variables (lower diagonal) and contemporaneous correlation (upper diagonal) from 1980 Q1 to 2009 Q2.

Correlation Analysis

Results

- Strong correlations between GDP and Housing investment (0.80) and Employment (0.72)
- Overall, contemporaneous correlation coefficients appear appear quite small in comparison of what could be expected
- Focus on cross-correlations

Cross-correlation with GDP



Cross-correlation Analysis

	GDP	Prices	Sales	Invest.	Employ.	Survey	Short	Long	Permits	Starts	
GDP		0.64	0.48	0.80	0.72	0.58	0.68	-0.52	0.58	0.62	0
Prices	-2		0.36	0.70	0.71	0.50	0.54	0.44	0.61	0.62	0
Sales	-3	-2		0.60	0.27	0.35	-0.37	-0.49	0.44	0.61	-
Investment	0	+1	+3		0.69	0.53	0.50	0.54	0.63	0.70	0
Employment	0	+2	+7	+1		0.55	0.56	0.64	0.52	0.47	0
Survey	-1	+1	+3	0	-2		0.37	-0.41	0.42	0.44	0
Short rate	0	+3	0	+1	0	+2		0.59	0.44	0.43	0
Long rate	-8	+3	-1	+3	0	-7	-1		-0.48	-0.47	-
Permits	-4	0	+1	-3	-4	-2	-5	+6		0.69	0
Starts	-1	0	+2	-1	-4	-1	-4	+5	+1		0
IPI	0	+2	-2	0	0	+2	-1	+8	+4	+3	

Table: Highest cross-correlation coefficients among all leads and lags (upper diagonal, lags in parenthesis) and leads/leags (lower diagonal), from 1980 Q1 to 2009 Q2. A negative number indicates that the series in row leads the series in column with an advance equal to this number, and conversely.

Cross-correlation with GDP

- A leading pattern in housing variables (Real prices, Sales, Permits and Starts)
- Residential investement is strongly related to the economic cycle, in coincident manner
- Employent and IPI in construction are coincident with GDP
- Short rate (3m): a positive correlation with a short delay (0-1 quarters)
- Long rate (10y): a negative correlation with lead of 2 years.

ARDL models

How to integrate dynamics into a mutivariate linear model?

Definition

$$Y_t = \alpha + \sum_{j=0}^{m} \beta_j' X_{t-j} + \sum_{j=1}^{p} \phi_j Y_{t-j} + \varepsilon_t, \qquad (1)$$

where:
X_t is the n-vector of explanatory variables $(X_{1t}, \ldots, X_{nt})'$,
m is the lag of the explanatory variables,
p is the AR order
ε_t strong WN,
for a given lag j, $\beta_j = (\beta_j^1, \ldots, \beta_j^n)'$ is the coefficient vector for explanatory variables of length n.

ARDL models

- The model specification is generally carried out using information criteria such as AIC or BIC.
- Note that m is not necessarily equal for all X_j, can be m_j, for $j = 1, \ldots, n$
- The $mn + p + 1$ parameters of the model can be estimated by ordinary least-squares

Forecasting based on ARDL

1. Iterative forecasting: Conditional forecasting of explanatory variables
2. Scenario forecasting: Judgemental forecasting of explanatory variables
3. Direct forecasting: a specific regression for each horizon h

Iterative forecasting based on ARDL

Assume $n = 1$ explanatory variable, $m = 1$, $p = 1$, $h = 1$:

$$Y_t = \alpha + \beta X_t + \phi Y_{t-1} + \varepsilon_t \tag{2}$$

$$\hat{Y}_t(1) = E(Y_{t+1}|I_t) = \hat{\alpha} + \hat{\beta} E(X_{t+1}|I_t) + \hat{\phi} Y_t \tag{3}$$

How to compute $E(x_{t+1}|I_t)$? Use of auxiliary models

Scenario forecasting based on ARDL

Example of 3 scenarii for $E(X_{t+1}|I_t)$:

1. negative growth of -2% (X^-_{t+1})
2. stability : 0% (X^0_{t+1})
3. positive growth of +2% (X^+_{t+1})

$$\hat{Y}^-_t(1) = \hat{\alpha} + \hat{\beta} X^-_{t+1} + \hat{\phi} Y_t \quad (4)$$

$$\hat{Y}^0_t(1) = \hat{\alpha} + \hat{\beta} X^0_{t+1} + \hat{\phi} Y_t \quad (5)$$

$$\hat{Y}^+_t(1) = \hat{\alpha} + \hat{\beta} X^+_{t+1} + \hat{\phi} Y_t \quad (6)$$

Direct forecasting based on ARDL

For any $h > 0$

$$Y_{t+h} = \alpha_h + \sum_{j=0}^{m} \beta'_{hj} X_{t-j} + \sum_{j=1}^{p} \phi_{hj} Y_{t-j} + \varepsilon_{t+h}, \qquad (7)$$

where, for a given lag j, $\beta_{hj} = (\beta^1_{hj}, \ldots, \beta^n_{hj})'$ is the coefficient vector for financial variables of length n.

The h-step-ahead forecast is thus given by

$$\hat{Y}_t(h) = \hat{\alpha}_h + \sum_{j=0}^{m} \hat{\beta}'_{hj} X_{t-j} + \sum_{j=1}^{p} \hat{\phi}_{hj} Y_{t-j+1}. \qquad (8)$$

References related GDP forecasting based on financial variables

US data:
Estrella and Hardouvelis (1991), Hamilton and Kim (2002), Estrella et al. (2003) or Giacomini and Rossi (2006).
Gilchrist and Zakrajsek (2012): a new credit spread index to predict US GDP growth.
Estrella and Mishkin (1997): usefulness of various term spreads and monetary variables for the US GDP
Euro area:
Andersson and d'Agostino (2008) use sectoral stock prices to predict the euro area GDP.
Duarte et al. (2005) spread between 10-year sovereign yeld and the 3-month interbank rate

Example: Taylor rule

Basic regression for central bank interest rates proposed by Taylor (1993):

$$r_t = \alpha + \beta g_t + \gamma(\pi_t - \pi*) + \varepsilon_t$$

where:
π_t is a measure of inflation (headline or core),
$\pi*$ is inflation targeting (assuming the CB is inflation targeter)
g_t a measure of slack, usually output gap
ε_t is a white noise.

$\beta = 0.5$ in Taylor (1993), but $\beta = 1$ often used
$\gamma = 1.5$

Example: Taylor rule

In this framework, α can be seen as the neutral nominal interest rate.

Basic regression becomes:

$$r_t = (r_t^* + \pi *) + \beta g_t + \gamma(\pi_t - \pi *) + \varepsilon_t$$

where:
r_t^* is the real nautral interest rate

Example: Taylor rule

Extended regression for central bank interest by accounting for persistence in the interest rate (inertial Taylor rule):

$$r_t = \rho r_{t-1} + \alpha + \beta g_t + \gamma(\pi_t - \pi *) + \varepsilon_t$$

where:
ρ controls the persistence (generally estimated around 0.85)

Example: Taylor rule

Variant of the extended regression:

$$r_t = \rho r_{t-1} + (1-\rho)\{\alpha + \beta g_t + \gamma(\pi_t - \pi*)\} + \varepsilon_t$$

See R codes for US Taylor rules estimation