Real GDP and Economic Indicators

LOOKING INTO THE FUTURE OF OUR ECONOMY

Contents

Eric Peterson
Liberty University
Undergraduate, Industrial & Systems Engineering
edpeterson2@liberty.edu

Real GDP and Economic Indicators

LOOKING INTO THE FUTURE OF OUR ECONOMY

This document will analyze the driving factors of United States Gross Domestic Product and predict the future value of GDP in the next quarter. Certain economic data can indicate how the GDP will change in the following quarter. Here we take a look into some of the most telling factors.

THE OBJECTIVE

The objective of this procedure is to model and predict the US GDP in the near future based on other economic ***indicators***. A history will be compiled and the effects of the indicators on its value will be analyzed. Then the current values of the indicators will be used to predict the next GDP value.

GATHERING THE DATA

Any number of indicators can be used, but five series have been chosen from www.economagic.com and www.data360.org. These are consumer price index, unemployment, prime rate, industrial production index, and inflation. We are comparing this with Real GDP in billions of 2000 dollars, published quarterly by the US Bureau of Economic Analysis.

In a spreadsheet, we will line up all this data in columns with headers for the different data sets.

Date	CPI	U	PR	IPI	Inf	GDP
1/1/1969	35.833	3.40	7.06	42.99	5.79	3,750.20
4/1/1969	36.433	3.43	7.74	43.18	5.67	3,760.90
7/1/1969	36.967	3.57	8.50	43.70	5.59	3,784.20
10/1/1969	37.500	3.57	8.50	43.43	6.63	3,766.30
1/1/1970	38.000	4.17	8.46	42.37	6.52	3,760.00
4/1/1970	38.633	4.77	8.00	42.14	5.34	3,767.10
7/1/1970	39.067	5.17	7.94	41.99	4.20	3,800.50
10/1/1970	39.600	5.83	7.23	41.08	6.26	3,759.80
1/1/1971	39.900	5.93	5.87	41.87	2.04	3,864.10
4/1/1971	40.333	5.90	5.41	42.25	5.11	3,885.90
7/1/1971	40.767	6.03	5.97	42.38	3.00	3,916.70
10/1/1971	40.967	5.93	5.64	43.37	2.97	3,927.90
1/1/1972	41.267	5.77	4.89	43.20	2.98	3,997.70
4/1/1972	41.600	5.70	5.00	46.07	2.93	4,092.10
7/1/1972	42.000	5.57	5.34	46.67	3.90	4,131.10
10/1/1972	42.400	5.37	5.76	48.27	3.86	4,198.70

The data available ranges from January (Q1) 1969 to October (Q4) 2008

The data I've collected dates back to 1969. However, if we use the most recent seven years, that will be a total of 28 quarters, which is an appropriate number for a ***multivariate regression***, the statistical analysis we will use.

CHOOSING A MODEL

Most multivariate regressions are linear, assuming the data takes the form

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n + \varepsilon$$

Where GDP Y is represented as a linear combination of all the different driver variables plus a noise variable ε. However, a more accurate ***variable elasticity*** model is

$$y = b_0 * e^{b_1x_1} * x_1^{b_2} * e^{b_3x_2} * x_2^{b_4} * \ldots * e^{b_{2n-1}x_n} * x_n^{b_{2n}}$$

However, this is very complex and nonlinear. If we take the natural log of both sides and simplify, the equation becomes

$$\ln y = b_0' + b_1x_1 + b_2 \ln x_1 + b_3x_2 + b_4 \ln x_2 + \cdots + b_{2n-1}x_n + b_{2n} \ln x_n$$

This form of a variable elasticity model is much easier estimated by linear multivariate means. Here you can see the actual model is of $\ln y$, but once we've forecasted the natural log of GDP, we can simply raise e to that power to have our actual estimate of GDP.

PERFORMING THE REGRESSION

To perform a regression on a spreadsheet, the data sets need to be organized into columns. In addition to our GDP column, we now need an "Ln GDP" column for our y values and columns for the natural logs of all of our x_i's. The GDP columns should be shifted so that they display the GDP one quarter *after* the indicators. In essence, we are lining up rows of nine inputs and one output. The inputs are the economic indicator values of each particular quarter and the output is the GDP one quarter later. The layout looks as so:

	x_1	x_2	x_3	x_4	x_5	ln x_1	ln x_2	ln x_3	ln x_4	y	ln y
Date	CPI	U	PR	IPI	Inf	Ln CPI	Ln U	Ln PR	Ln IPI	**GDP in 1 Qtr**	**ln GDP**
1/1/2002	177.900	5.70	4.75	98.69	2.51	5.1812	1.7405	1.5581	4.5920	10,031.60	9.213495
4/1/2002	179.833	5.83	4.75	100.11	2.51	5.1920	1.7630	1.5581	4.6063	10,090.70	9.219369
7/1/2002	180.600	5.73	4.75	100.64	2.70	5.1963	1.7457	1.5581	4.6116	10,095.80	9.219875
10/1/2002	181.167	5.87	4.45	100.55	2.23	5.1994	1.7699	1.4929	4.6107	10,126.00	9.222862
1/1/2003	183.000	5.87	4.25	101.28	4.72	5.2095	1.7699	1.4469	4.6179	10,212.70	9.231387
4/1/2003	183.667	6.13	4.24	100.52	-1.70	5.2131	1.8132	1.4446	4.6103	10,398.70	9.249436
7/1/2003	184.567	6.13	4.00	101.15	4.44	5.2180	1.8132	1.3863	4.6166	10,467.00	9.255983
10/1/2003	184.600	5.83	4.00	102.18	0.88	5.2182	1.7630	1.3863	4.6268	10,543.60	9.263274
1/1/2004	186.267	5.70	4.00	102.89	3.50	5.2272	1.7405	1.3863	4.6337	10,634.20	9.271831

These are our x-values for the regression.

Select these y-values.

In the data analysis tools (I'm using MS Excel 2007), we now perform a regression on our data, selecting the rows of indicators as the x-input and the column of ln GDP as the y-input. The output will contain the coefficients of our model, as well as other statistical values.

Notice that there is no $\ln x_5$ column because inflation contains negative values, which are not in the domain of the ln function. For simplicity I have omitted this column, although it would also be possible to make a column for $\ln(x_5 + c)$, where c is some constant added to all the inflation values to make them all positive.

READING THE OUTPUT

The regression output will be a box that looks like this:

SUMMARY OUTPUT						
Regression Statistics			n	28		
Multiple R	0.98954		Ybar	9.299		
R Square	0.97918		n*Ybar^2	2421		
Adjusted R Square	0.97026		sumsqY	2508		
Standard Error	0.01992		k	0.96549		
Observations	31		R(U)	0.99975		
ANOVA						
	df	*SS*	*MS*	*F*	*Significance F*	
Regression	9	0.391884	0.043543	109.74151	1.30015E-15	
Residual	21	0.008332	0.000397			
Total	30	0.400217				
	Coefficients	*Standard Error*	*t Stat*	*P-value*	*Lower 95%*	*Upper 95%*
Intercept	91.1092	23.7398	3.8378	0.0010	41.7396	140.4787
CPI	0.043002	0.0338	1.2712	0.2176	-0.0273	0.1134
Unemp	0.089779	0.1752	0.5124	0.6137	-0.2746	0.4542
PR	0.137302	0.0354	3.8764	0.0009	0.0636	0.2110
IPI	0.142588	0.1018	1.4004	0.1760	-0.0692	0.3543
Inf	-0.001782	0.0011	-1.5628	0.1331	-0.0042	0.0006
Ln CPI	-7.50330	7.0238	-1.0683	0.2975	-22.1102	7.1036
Ln Unemp	-0.57511	0.9886	-0.5817	0.5669	-2.6311	1.4808
Ln PR	-0.81405	0.1943	-4.1904	0.0004	-1.2181	-0.4101
Ln IPI	-13.85876	11.1129	-1.2471	0.2261	-36.9693	9.2518

This is the uncorrected R-square that we compute.

The b_i coefficients are listed in the "Coefficients" column, and many other statistics are included as well. One statistic that is helpful to compute is the uncorrected r-square. To do this, we first compute the k value, which is the number of y-values, 28, times the square of the average y-value, divided by the sum of all the squared y-values. The R(U) square is the R square times $1 - k$ plus the k value. Ours is 0.99975, which means our model explains 99.975% of the data in our history.

Our specific model for GDP is

$$\ln GDP = 91.1 + 0.043 * CPI + 0.090 * Unemp + 0.137 * Prime\ Rate + 0.143 * IPI - 0.00178 * Inflation - 7.5 * \ln CPI - 0.575 * \ln Unemp - 0.814 * \ln PR - 13.859 * \ln IPI$$

And $$GDP = e^{\ln GDP}$$

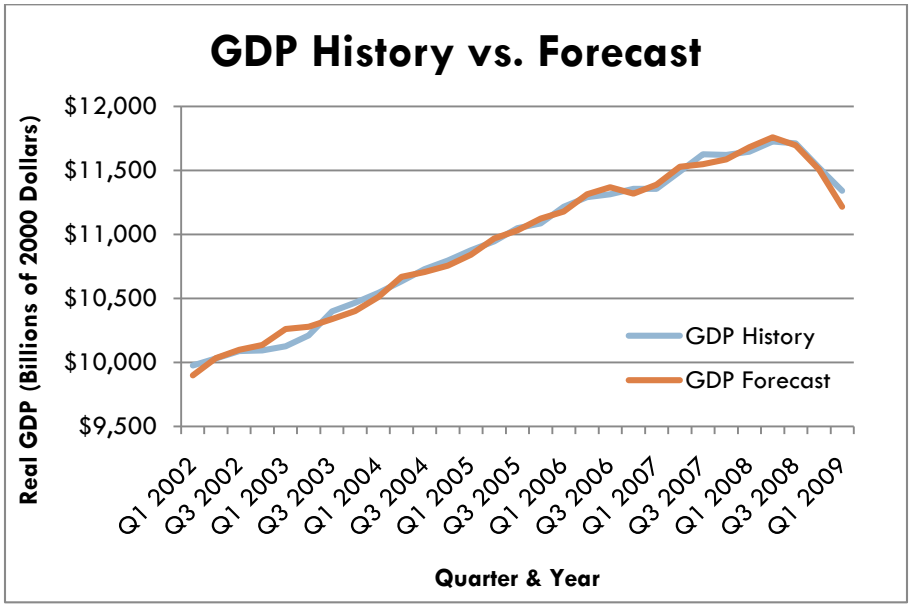
APPLYING THE MODEL AND FORECAST

Now that we've calculated the coefficients to our model, we can use our historical drivers to make a model of GDP and use the drivers that we know now to estimate the current GDP.

Date	GDP	ln GDP	Mod ln	Model	|A%E|
4/1/2006	11,291.70	9.3318	9.387825	11942.1	5.76%
7/1/2006	12,965.90	9.4701	9.431378	12473.7	3.80%
10/1/2006	13,060.70	9.4774	9.464055	12888.04	1.32%
1/1/2007	13,099.90	9.4804	9.460402	12841.04	1.98%
4/1/2007	13,204.00	9.4883	9.473027	13004.2	1.51%
7/1/2007	13,321.10	9.4971	9.510416	13499.6	1.34%
10/1/2007	13,391.20	9.5024	9.51752	13595.86	1.53%
1/1/2008	13,366.90	9.5005	9.500241	13362.94	0.03%
4/1/2008	13,415.30	9.5042	9.496445	13312.32	0.77%
7/1/2008	13,324.60	9.4974	9.503823	13410.9	0.65%
10/1/2008	13,141.90	9.4836	9.48456	13155.04	0.10%
1/1/2009	12,925.40	9.4669	9.45791	12809.09	0.90%
4/1/2009	12,892.40	9.4644	9.469185	12954.32	0.48%
7/1/2009			9.516754	**13585.44**	

Since we know the current values of all our drivers, we can calculate the GDP for next quarter. Based on our data, the model says first quarter Real GDP for the 3rd Quarter of 2009 will be 13,585 billion 2000 dollars, a start of a recovery. On April 29, 2009 the Bureau of Economic Analysis published the Q1 2009 GDP as 11,340.9 billion 2000 dollars. All these drivers were available well before the release and predicted it within 98.9%.

The column labeled |A%E| is the absolute percent error, which is the absolute value of the difference between the model and historical, divided by the historical value. Most are all well within 98%. The |AA%E|, average absolute percent error, is 1.14%. This is a very accurate model.



SUMMARY

This forecast can be done in a few easy steps. First, choose the value you want to forecast. This could be GDP, the price of a certain product, total sales in a company, or any other desired value. Next, choose a handful of indicators that you think might be influencing that value. The indicators I've listed above are only a small amount of all the economic indicators out there. For business, indicators can be marketing budgets, capital spending, depreciation, total assets, stockholder's equity, and many more.

Once the target and indicators have been chosen, gather their historical values across a few dozen intervals and shift the target values forward as far as you want to predict. In essence, gathering the data by months and shifting the target two months forward will be like asking, "How do these values on such-and-such a month correlate with the desired value two months later?"

Perform the regression. The large box of all the indicators will be the x's, and the target column will be the y's. Check the accuracy of the regression – it is usually given with the output.

Use the coefficients given in the regression output to construct a model of the target value. It will always be of the basic form:

$$y = b_0 + b_1x_1 + b_2x_2 + \cdots + b_nx_n$$

Now use this formula to model the values in your history. If the model values are similar to the historical values, the model does a good job explaining the data. Now use that same formula and enter the most current values of the indicators, and the result will be the forecasted future value. When the parameters are chosen correctly, this will be a very accurate forecast.