Multiple Regression on the Example of the Presidential Elections in the United States during the Period from 1916 to 2000

Taenvat MM*

Vladivostok Postgraduate of the Department of State and Municipal Management Institute of Law and Governance, Vladivostok State University Economics and Service, Russia

Abstract

Systematization of knowledge about multiple regression: the withdrawal of the regression equation 1) Through a system of normal equations and 2) Matrix method; solution 1) Using packet analysis EXCEL, 2) Matrix calculator and "matrix arithmetic"; 3) Solution in MATLAB as the normal equations, and directly – matrix method; controversy about the statistical reliability - University of Vladivostok against the University of Indiana; graphical and tabular analysis of residues; concluded by differentiating the normal equations; Study regression with z-statistics and graphs standard normal probability.

Keywords: Multiple regression-systematization; Controversial statistical reliability; Regression equation; MATLAB; EXCEL; Matrix calculator; Analysis of residues; Concluded by differentiating the normal equations; z-Statistics

Introduction

Multiple regressions are studied in Russia and abroad. So Chetyrkin [1] describe a method of constructing it manually. Winston [2,3] described the quickest way to build it in the program Microsoft Office Excel, and taken them an example of submission of multiple regression, as well as details of its interpretation of some of the best works on the statistics. Sachs [4] not a lot of concerns of multiple regression, but look at how he described the manual methods of investigation regression really understand what the German precision [5].

But outside attention the authors were:

1. Systematization of knowledge about the multiple regressions;
2. Construction of the system of normal equations for mathematical models with more than two explanatory variables; conclusion of these equations is considered redundant because, firstly, the need to build an auxiliary table, and secondly, the solution of these equations requires a lot of work if it is to perform manually. But if you examine the sum of squared errors, standard errors of prediction and covariance, and you’ll need. A solution of large systems is now easier MATLAB. Plus the consideration of such a system shows the interaction of variables;
3. Conclusion of normal equations of the functional equation by differentiation;
4. The decision referred to in MATLAB as the normal equations, and direct-matrix method;
5. The decision of using a calculator and matrix "matrix arithmetic";
6. Graphical analysis of residues;
7. Investigation of multiple regressions with z - statistics and graphs standard normal probability. The regression statistics - the most difficult section of statistics. In turn, the multiple regression - one of the most difficult in the regression statistics. The formula is as follows:

\[ y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \ldots + \beta_n x_{ni} + \epsilon_i \]  

(1)

Chetyrkin [1] wrote: "If selected as the independent variable is the dominant factor, respectively, the corresponding pair regression adequately describes the mechanism of causation. The most common change of y due to the influence of several factors (sometimes acting in opposite directions). In these cases, naturally the desire to enter some explanatory variables. This is called multiple regressions. Multiple regression equation to better explain the behavior of the dependent variable than steam regression, in addition, it makes it possible to compare the effectiveness of various factors".

Multiple Regressions

Wayne Winston – Indiana University professor who advises the company Ford Motor corporation General Motors, Intel, Microsoft, Proctor and Gamble, the US Army, US Department of Defense and other organizations, a graduate of Yale University with a Ph.D. and the Faculty of Mathematics at MIT, who among universities in the world by Times Higher Education World Reputation Rankings 2015 took 4th place, refers to a book economist Roy Fair [6] that the economy has a major impact on the results of the presidential elections. How can I predict U.S. presidential elections?

Presidential advisor James Carville said "It’s the economy" when asked about which factors drive presidential elections. Yale economist Roy Fair showed that Carville was right in thinking that the state of the economy has a large influence on the results of presidential elections. Fair’s dependent variable for each election (1916–2000) was the % of the two party vote (ignoring votes received by third party candidates) that went to the incumbent party. He tried to predict the incumbent party’s % of the two party votes by using the following independent variables:

1) Party in power. In our data, we use a 1 to denote that the
Republican Party is in power and a 0 to denote that the Democratic Party is in power.

2) % growth in GNP during the first nine months of the election year.

3) Absolute value of the inflation rate during the first nine months of the election year. We use the absolute value because either a positive or a negative inflation rate is bad.

4) Number of quarters during the last four years in which economic growth has been strong. Strong economic growth is defined as growth at an annual level of 3.2% or more.

5) Time incumbent party has been in office. Fair used 0 to denote one term in office, 1 for two terms, 1.25 for three terms, 1.5 for four terms, and 1.75 for at least five terms. This definition implies that each term after the first term in office has less influence on the election results than the first term in office.

Is the election during wartime? The elections in 1920 (World War I), 1944 (World War II), and 1948 (World War II was still underway in 1945) were defined as wartime elections. Elections held during the Vietnam War were not considered wartime elections. During wartime years, the variables related to quarters of good growth and inflation was deemed irrelevant and was set to 0.

6) Is the current president running for re-election? If so, this variable is set to 1; otherwise, this variable is set to 0. In 1976, Gerald Ford was not considered a president running for re-election because he was not elected as either president or vice-president.

I've attempted to use the data from the elections in 1916 through 1996 to develop a multiple regression equation that can be used to forecast future presidential elections. I saved the infamous 2000 election as a «validation point».

In Table 1, you can see that the p-value for each independent variable is much less than .15, which indicates that each of our independent variables is helpful in predicting presidential elections. We can predict elections using an equation…:

\[
\text{Projected percentage of votes in the presidential election} = 45.53 + 0.70 \times \text{growth GNP} - 0.71 \times \text{abs.inflation} + 0.90 \times \text{quarter growth GNP} - 3.33 \times \text{terms of power} + 5.66 \times \text{Republican} + 4.71 \times \text{war} + 3.99 \times \text{President of a new term}
\]

The coefficients of the independent variables can be interpreted as follows

(After adjusting for all other independent variables used in equation 2):

1. A 1% increase in the annual GNP growth rate during an election year is worth .7% to the incumbent party.

2. A 1% deviation from the ideal (0% inflation) costs the incumbent party .71% of the vote.

3. Every good quarter of growth during an incumbent’s term increases his (maybe her someday soon!) vote by .90%.

4. Relative to having one term in office, the second term in office decreases the incumbent’s vote by 3.33%, and each later term decreases the incumbent’s vote by .25*(3.33%) = .83%.

5. A Republican has a 5.66% edge over a Democrat.

<table>
<thead>
<tr>
<th>Applicants</th>
<th>Year</th>
<th>The share of government</th>
<th>Ruling party</th>
<th>The growth rate of GNP in the election year</th>
<th>The absolute value of the rate of inflation in an election year</th>
<th>The number of quarters with GNP growth &gt;3.2%</th>
<th>The number of time spent by the ruling party in power</th>
<th>The ruling party 1 = republics Kanzen</th>
<th>War</th>
<th>Pull out whether the President for the next term?</th>
<th>Forecast</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ullikson - Hyyuz</td>
<td>1916</td>
<td>51.7</td>
<td>D</td>
<td>2.2</td>
<td>4.3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>50,71565364</td>
</tr>
<tr>
<td>Harding - Cox</td>
<td>1920</td>
<td>36.1</td>
<td>D</td>
<td>-11.5</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>38,86253503</td>
</tr>
<tr>
<td>Coolidge - Davis</td>
<td>1924</td>
<td>58.2</td>
<td>R</td>
<td>-3.9</td>
<td>5.2</td>
<td>10</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>57.78816261</td>
</tr>
<tr>
<td>Hoover - Smith</td>
<td>1928</td>
<td>58.8</td>
<td>R</td>
<td>4.6</td>
<td>0.2</td>
<td>7</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>57.28141426</td>
</tr>
<tr>
<td>Guver - FDR</td>
<td>1932</td>
<td>40.8</td>
<td>R</td>
<td>-14.9</td>
<td>7.1</td>
<td>4</td>
<td>1.25</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>39,14128587</td>
</tr>
<tr>
<td>FDR - Lendon</td>
<td>1936</td>
<td>62.5</td>
<td>D</td>
<td>11.9</td>
<td>2.4</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>64,29354145</td>
</tr>
<tr>
<td>FDR - Ullik</td>
<td>1940</td>
<td>55</td>
<td>D</td>
<td>3.7</td>
<td>0</td>
<td>8</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>56,019</td>
</tr>
<tr>
<td>FDR - Dewey</td>
<td>1944</td>
<td>53.8</td>
<td>D</td>
<td>4.1</td>
<td>0</td>
<td>0</td>
<td>1.25</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>52,94119656</td>
</tr>
<tr>
<td>Truman - Dudy</td>
<td>1948</td>
<td>52.4</td>
<td>D</td>
<td>1.8</td>
<td>0</td>
<td>0</td>
<td>1.5</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>50,49626842</td>
</tr>
<tr>
<td>Ike - Stevenson</td>
<td>1952</td>
<td>44.6</td>
<td>D</td>
<td>0.6</td>
<td>2.3</td>
<td>6</td>
<td>1.75</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>43,91467348</td>
</tr>
<tr>
<td>Ike - Stevenson</td>
<td>1956</td>
<td>57.8</td>
<td>R</td>
<td>-1.5</td>
<td>1.9</td>
<td>5</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>57,30571419</td>
</tr>
<tr>
<td>Kennedy - Nixon</td>
<td>1960</td>
<td>49.9</td>
<td>R</td>
<td>0.1</td>
<td>1.9</td>
<td>5</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>51,10831446</td>
</tr>
<tr>
<td>Johnson - Goldwater</td>
<td>1964</td>
<td>61.3</td>
<td>D</td>
<td>5.1</td>
<td>1.2</td>
<td>10</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>61,29837597</td>
</tr>
<tr>
<td>Nixon - Humphrey</td>
<td>1968</td>
<td>49.6</td>
<td>R</td>
<td>4.8</td>
<td>3.2</td>
<td>7</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>49,62176954</td>
</tr>
<tr>
<td>Nixon - MakGovern</td>
<td>1972</td>
<td>61.8</td>
<td>R</td>
<td>6.3</td>
<td>4.8</td>
<td>4</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>59,7995525</td>
</tr>
<tr>
<td>Ford - Carter</td>
<td>1976</td>
<td>48.9</td>
<td>R</td>
<td>3.7</td>
<td>7.7</td>
<td>4</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>48,59420464</td>
</tr>
<tr>
<td>Carter - Reagan</td>
<td>1980</td>
<td>44.7</td>
<td>D</td>
<td>-3.8</td>
<td>8.1</td>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>45,61537819</td>
</tr>
<tr>
<td>Reagan - Mondale</td>
<td>1984</td>
<td>59.2</td>
<td>R</td>
<td>5.4</td>
<td>5.4</td>
<td>7</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>61,45627149</td>
</tr>
<tr>
<td>Bush - Dukakis</td>
<td>1988</td>
<td>53.9</td>
<td>R</td>
<td>2.1</td>
<td>3.3</td>
<td>6</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>52,41708877</td>
</tr>
<tr>
<td>Bush - Clinton</td>
<td>1992</td>
<td>46.5</td>
<td>R</td>
<td>2.3</td>
<td>3.7</td>
<td>1</td>
<td>1.25</td>
<td>1</td>
<td>0</td>
<td>1</td>
<td>50,89799193</td>
</tr>
<tr>
<td>Clinton - Dole</td>
<td>1996</td>
<td>54.7</td>
<td>D</td>
<td>2.9</td>
<td>2.3</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>52,63060779</td>
</tr>
<tr>
<td>Gore - Bush</td>
<td>2000</td>
<td>50.3</td>
<td>D</td>
<td>2.2</td>
<td>1.7</td>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>52,20336851</td>
</tr>
</tbody>
</table>

Table 1: Data on presidential elections.
6. A wartime incumbent president has a 4.71% edge over his opponent.

7. A sitting president running for re-election has a 3.99% edge over his opponent.

We find that 94% of the variation in the % received by an incumbent in a presidential election is explained by our independent variables. We have made no mention whether the candidates are "good or bad" candidates (Table 1).

The author of the article draws attention to the fact that the fall in the growth rate of GDP in the election year, almost always leads to a change of the ruling party (Table 1).

First, we decide through the normal equations.

\[
\begin{align*}
\sum_{i=1}^{n} a_i x_i & = \sum_{i=1}^{n} y_i, \\
\sum_{i=1}^{n} a_i x_i^2 & = \sum_{i=1}^{n} x_i y_i, \\
\sum_{i=1}^{n} a_i x_i^3 & = \sum_{i=1}^{n} x_i^2 y_i, \\
\sum_{i=1}^{n} a_i x_i^4 & = \sum_{i=1}^{n} x_i^3 y_i, \\
\sum_{i=1}^{n} a_i x_i^5 & = \sum_{i=1}^{n} x_i^4 y_i, \\
\sum_{i=1}^{n} a_i x_i^6 & = \sum_{i=1}^{n} x_i^5 y_i.
\end{align*}
\]

The output of this system by differentiation:

\[
E(a, b, c, d, e, h, p, q) = \sum_{i=1}^{n} u_i, \quad \text{if} \quad u = ax + cx + dx + ex + hx + px + qx - y.
\]

Since we have eight variables a, b, c, d, e, h, p, and q – we take the partial derivatives. Fix b, c, d, e, h, p, and q and differentiate E (a, b, c, d, e, h, p, q) on a. Obtain

\[
\frac{\partial E}{\partial a} = \sum_{i=1}^{n} u_i, \quad \text{if} \quad u = ax + cx + dx + ex + hx + px + qx - y.
\]

We fix a, c, d, e, h, p and q and differentiate E (a, b, c, d, e, h, p, q) by b. Receive

\[
E(a, b, c, d, e, h, p, q) = \sum_{i=1}^{n} u_i, \quad \text{if} \quad u = ax + cx + dx + ex + hx + px + qx - y.
\]

We fix a, b, c, d, e, h, p, and q and differentiate E (a, b, c, d, e, h, p, q) by c. Receive

\[
E(a, b, c, d, e, h, p, q) = \sum_{i=1}^{n} u_i, \quad \text{if} \quad u = ax + cx + dx + ex + hx + px + qx - y.
\]

We fix a, b, c, d, e, h, p, and q and differentiate E (a, b, c, d, e, h, p, q) by d. Receive

\[
E(a, b, c, d, e, h, p, q) = \sum_{i=1}^{n} u_i, \quad \text{if} \quad u = ax + cx + dx + ex + hx + px + qx - y.
\]
We fix \(a, b, c, d, h, p\) and \(q\) and differentiate \(E(a, b, c, d, e, h, p, q)\) by \(e\). Receive

\[
E(a, b, c, d, e, h, p, q) = \sum_{k=1}^{n} a_k' \text{ if } u = a + b x_1 + c x_2 + d x_3 + e x_4 + h x_5 + p x_6 + q x_7 - y.
\]

Here \[
\frac{\partial E(a, b, c, d, e, h, p, q)}{\partial e} = \sum_{k=1}^{n} \frac{\partial a_k'}{\partial e} = \sum_{k=1}^{n} 2 a_k x_k.
\]

Repectively, \[
\frac{\partial E(a, b, c, d, e, h, p, q)}{\partial e} = \sum_{k=1}^{n} \frac{\partial E(a, b, c, d, e, h, p, q)}{\partial e} = \sum_{k=1}^{n} 2 a_k x_k.
\]

We fix \(a, b, c, d, e, h, p\) and \(q\) and differentiate \(E(a, b, c, d, e, h, p, q)\) by \(h\). Receive

\[
E(a, b, c, d, e, h, p, q) = \sum_{k=1}^{n} a_k' \text{ if } u = a + b x_1 + c x_2 + d x_3 + e x_4 + h x_5 + p x_6 + q x_7 - y.
\]

Here \[
\frac{\partial E(a, b, c, d, e, h, p, q)}{\partial h} = \sum_{k=1}^{n} \frac{\partial a_k'}{\partial h} = \sum_{k=1}^{n} 2 a_k x_k.
\]

Repectively, \[
\frac{\partial E(a, b, c, d, e, h, p, q)}{\partial h} = \sum_{k=1}^{n} \frac{\partial E(a, b, c, d, e, h, p, q)}{\partial h} = \sum_{k=1}^{n} 2 a_k x_k.
\]

To find the coefficients of the system of equations we construct auxiliary (Tables 2a and 2b).

Finally got a model that reflects the policy of the United States for nearly a century:

\[
A = \begin{bmatrix}
21a_1 + 26a_2 + 65a_3 + 104a_4 + 14a_5 + 10a_6 + 3a_7 + 14a_8,
26a_1 + 7375a_2 + 12a_3 + 29a_4 + 240a_5 + 0.625a_6 + 4.2a_7 + 5.6a_8 + 21.6a_9,
12a_1 + 322a_2 + 332a_3 + 342a_4 + 32a_5 + 41a_6 + 46a_7 = 3346.5,
10a_1 + 4a_2 + 5a_3 + 6a_4 + 10a_5 + 6a_6 + 8a_7 = 558.5,
3a_1 - 5a_2 - 3.5a_3 + 3a_4 + 2a_5 = 1.423,
14a_1 + 21a_2 + 46a_3 + 69a_4 + 6.25a_5 + 6a_6 + 2a_7 + 14a_8 = 760.4.
\end{bmatrix}
\]

To solve the program will use MATLAB. First make sure that the matrix \(A\) – non-singular, showing that det \((A) \neq 0\).

\[
\text{det}(A) = 1.8924e + 10
\]

Decision systems \(AX = B\) represented in the form \(X = A^{-1}B\).
Table 2a: Auxiliary table.

| x1*x2 | x2*x3 | x3*x4 | x4*x5 | x5*x6 | y | y*x1 | y*x2 | y*x3 | y*x4 | y*x5 | y*x6 | y*x7 |
|-------|-------|-------|-------|-------|---|------|------|------|------|------|------|------|------|
| 0     | 2.2   | 4.3   | 3     | 0     | 0  | 6.6  | 51.7 | 113.74 | 222.31 | 155.1 | 0   | 0   | 51.7 |
| 0     | -11.5 | 0     | 0     | 0     | 0  | 36.1 | -415.15 | 0   | 36.1 | 0   | 361 | 0   | 0   |
| -3.9  | 0     | -3.9  | 5.2   | 10    | 1  | 5.2  | -39  | 58.2 | -226.98 | 302.64 | 582 | 0   | 582 |
| 4.6   | 0     | 0     | 0     | 0     | 0  | 32.2 | 58.2 | 270.48 | 117.6 | 411.6 | 58.8 | 0   | 0   |
| -14.9 | 0     | -14.9 | 7.1   | 4     | 1  | 7.1  | -59.6 | 40.8 | -607.92 | 289.68 | 163.2 | 51  | 40.8 |
| 0     | 11.9  | 2.4   | 9     | 0     | 0  | 107.1 | 62.5 | 743.75 | 150   | 562.5 | 0   | 0   | 62.5 |
| 0     | 3.7   | 0     | 8     | 0     | 0  | 29.6 | 55   | 203.5 | 0     | 440  | 55  | 0   | 55  |
| 0     | 4.1   | 4.1   | 0     | 0     | 0  | 53.8 | 220.58 | 0   | 67.25 | 0   | 53.8 | 0   | 53.8 |
| 0     | 1.8   | 1.8   | 0     | 0     | 0  | 52.4 | 94.32 | 0   | 78.6 | 0   | 52.4 | 0   | 52.4 |
| 0     | 0     | 0     | 0     | 0     | 0  | 3.6  | 44.6 | 26.76 | 102.58 | 267.6 | 78.05 | 0   | 0   |
| -1.5  | 0     | -1.5  | 1.9   | 5     | 1   | 1.9  | -7.5 | 57.8 | -86.7 | 109.82 | 289  | 0   | 57.8 |
| 0.1   | 0     | 0     | 0     | 0     | 0  | 1.9  | 0.5  | 49.9 | 4.99 | 94.81 | 249.5 | 49.9 | 0   |
| 0     | 0     | 5.1   | 1.2   | 10    | 0  | 51   | 61.3 | 312.63 | 73.56 | 613  | 0   | 0   | 61.3 |
| 0     | 0     | 0     | 0     | 0     | 0  | 33.6 | 49.6 | 238.08 | 158.72 | 347.2 | 49.6 | 0   | 0   |
| 6.3   | 0     | 6.3   | 4.8   | 4     | 1   | 4.8  | 25.2 | 61.8 | 389.34 | 296.64 | 247.2 | 0   | 61.8 |
| 3.7   | 0     | 0     | 0     | 0     | 0  | 7.7  | 14.8 | 48.9 | 180.93 | 376.53 | 195.6 | 48.9 | 0   |
| 0     | 0     | -3.8  | 8.1   | 5     | 0  | 0    | -19 | 44.7 | -169.86 | 362.07 | 223.5 | 0   | 0   |
| 5.4   | 0     | 5.4   | 7     | 5     | 4  | 37.8 | 59.2 | 319.68 | 319.68 | 414.4 | 0   | 59.2 | 0   |
| 2.1   | 0     | 0     | 0     | 0     | 3  | 0    | 12.6 | 53.9 | 113.19 | 177.87 | 323.4 | 53.9 | 0   |
| 2.3   | 0     | 2.3   | 3     | 1     | 1   | 3.7  | 2.3  | 46.5 | 106.95 | 172.05 | 46.5 | 58.125 | 46.5 |
| 0     | 0     | 2.9   | 2.3   | 3     | 0  | 0    | 8.7  | 54.7 | 158.63 | 125.81 | 164.1 | 0   | 0   |
| 4.2   | -5.6  | 21.6  | 46.4  | 69    | 6   | 41.2 | 0    | 240.5 | 1102.2 | 1990.94 | 3346.53 | 5695.4 | 685.225 |

Table 2b: Auxiliary table.
X = inv(A) \* [1102.200 1990.94 3346.53 5695.400 685.225 535.810 142.300 760.400]

\[
X = \begin{bmatrix}
1 & x_1 & x_2 & \cdots & x_n \\
1 & x_{11} & x_{12} & \cdots & x_{1n} \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
1 & x_{m1} & x_{m2} & \cdots & x_{mn}
\end{bmatrix}
\]

\[
X' = \begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_m
\end{bmatrix}
\]

\[
X'X = \begin{bmatrix}
\sum y_i \\
\sum y_i x_{i1} \\
\vdots \\
\sum y_i x_{in}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\sum y_i \\
\sum y_i x_{i1} \\
\vdots \\
\sum y_i x_{in}
\end{bmatrix}
\]

\[
\begin{bmatrix}
\begin{bmatrix}
\sum y_i \\
\sum y_i x_{i1} \\
\vdots \\
\sum y_i x_{in}
\end{bmatrix}
\end{bmatrix}
\]

\[
a = (X'X)^{-1} X' y \quad y = X a + e
\]

We test the solution, making sure that AX = B.

B = A \* X

\[
B = \begin{bmatrix}
1.0000 \\
1.0000 \\
\vdots \\
1.0000
\end{bmatrix}
\]

We equate the result to zero. After that, we can easily find the system of normal equations that in matrix form is written as

\[
X'X a = X'y
\]

The sum of squared deviations obtained as follows:

\[
Q = \sum e_i^2 = (y - Xa)^T (y - Xa) = y' y - a' X' y - a' X' X a + a' X' X a
\]

Here and below T denotes the transpose of a vector or matrix. So how

\[
a' X' y = y' X a,
\]

then

\[
Q = y' y - 2a' X' y + a' X' X a
\]

Differentiating Q for a and we get

\[
a' = \frac{dQ}{da} = 2 X' y + 2 a' X' X
\]

Now to solve MATLAB.

\[
X = \begin{bmatrix}
1 & 2.2 & 4.3 & 3 & 0 & 0 & 0 & 1 \\
1 & -11.5 & 0 & 0 & 1 & 0 & 1 & 0 \\
1 & -3.9 & 5.2 & 10 & 0 & 1 & 0 & 1 \\
1 & 4.6 & 0.2 & 7 & 1 & 1 & 0 & 1 \\
1 & -14.9 & 7.1 & 4 & 1.25 & 1 & 0 & 1 \\
1 & 3.7 & 0 & 8 & 1 & 0 & 0 & 1 \\
1 & 1.8 & 0 & 0 & 1.5 & 0 & 1 & 0 \\
1 & 0.6 & 2.3 & 6 & 1.75 & 0 & 0 & 0 \\
1 & -1.5 & 1.9 & 5 & 0 & 1 & 1 & 0 \\
1 & 0.1 & 1.9 & 5 & 1 & 1 & 0 & 0 \\
1 & 5.1 & 1.2 & 10 & 0 & 0 & 0 & 1 \\
1 & 4.8 & 3.2 & 7 & 1 & 0 & 0 & 0 \\
1 & 6.3 & 4.8 & 4 & 0 & 1 & 0 & 1 \\
1 & 3.7 & 7.7 & 4 & 1 & 1 & 0 & 0 \\
1 & -3.8 & 8.1 & 5 & 0 & 0 & 0 & 1 \\
1 & 5.4 & 5.4 & 7 & 0 & 1 & 1 & 0 \\
1 & 2.1 & 3.3 & 6 & 1 & 1 & 0 & 0 \\
1 & 2.3 & 3.7 & 1 & 1.25 & 1 & 0 & 1 \\
1 & 2.9 & 2.3 & 3 & 0 & 0 & 0 & 1 \\
\end{bmatrix}
\]

\[
X' \quad \text{ans} = \begin{bmatrix}
1.0000 \\
1.0000 \\
\vdots \\
1.0000
\end{bmatrix}
\]

Formula solutions method takes the form:
### Table 1: Example Data

<table>
<thead>
<tr>
<th>x1</th>
<th>x2</th>
<th>x3</th>
<th>x4</th>
<th>y</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.7000</td>
<td>-3.8000</td>
<td>5.4000</td>
<td>2.1000</td>
<td>2.3000</td>
</tr>
<tr>
<td>4.3000</td>
<td>0.5200</td>
<td>0.2000</td>
<td>7.1000</td>
<td>2.4000</td>
</tr>
<tr>
<td>5.4000</td>
<td>3.3000</td>
<td>3.7000</td>
<td>2.3000</td>
<td>3.0000</td>
</tr>
<tr>
<td>6.0000</td>
<td>6.0000</td>
<td>1.0000</td>
<td>3.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>1.7500</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>0 0 1.0000</td>
<td>1.0000</td>
<td>0 1.0000</td>
<td>1.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>1.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>0 0 0 0 0 0 0 0 0 0 0 1.0000</td>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td></td>
</tr>
<tr>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
</tr>
<tr>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td></td>
</tr>
<tr>
<td>1.0000</td>
<td>0 1.0000</td>
<td>0 1.0000</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Table 2: Calculated Values

<table>
<thead>
<tr>
<th>X'X ans</th>
<th>X'y ans</th>
<th>(X'X)^{-1}*(X'y) ans</th>
</tr>
</thead>
<tbody>
<tr>
<td>21.0000</td>
<td>26.0000</td>
<td>65.0000</td>
</tr>
<tr>
<td>737.5800</td>
<td>12.2900</td>
<td>240.5000</td>
</tr>
<tr>
<td>12.2900</td>
<td>332.9000</td>
<td>342.2000</td>
</tr>
<tr>
<td>342.2000</td>
<td>706.0000</td>
<td>53.7500</td>
</tr>
<tr>
<td>332.9000</td>
<td>342.2000</td>
<td>53.7500</td>
</tr>
<tr>
<td>342.2000</td>
<td>53.7500</td>
<td>53.7500</td>
</tr>
<tr>
<td>706.0000</td>
<td>53.7500</td>
<td>53.7500</td>
</tr>
<tr>
<td>420.0000</td>
<td>342.2000</td>
<td>53.7500</td>
</tr>
<tr>
<td>120.0000</td>
<td>332.9000</td>
<td>342.2000</td>
</tr>
<tr>
<td>420.0000</td>
<td>342.2000</td>
<td>53.7500</td>
</tr>
</tbody>
</table>

We can solve the problem and using the calculator matrices [7] and "matrix arithmetic". Enter the original data into the calculator (Figure 1). We find the transposed matrix (Figure 2). Go to the window "matrix arithmetic" (Figure 3). We get work and transpose of the original matrix (Figure 4). We find the inverse matrix (Figures 5 and 6). In the same window, find the product of the transposed matrix of the vector and the column y. And by multiplying the inverse matrix that happened, we obtain a matrix of unknown coefficients (Figures 7 and 8) (Tables 3-7).

1. Since 29.4043 ≥ 2.832097502 and 5.48E-07 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;
2. Since 7.458777 ≥ 2.160396562, a 4.77E-06 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;
Figure 3: Window matrix arithmetic.

Figure 4: The result of matrix multiplication.

Figure 5: Calculator matrices.

Figure 6: The inverse matrix.

Figure 7: Decision.

Figure 8: Fair R. predicting presidential elections and other things.

Figure 9: US presidential elections.

Regression Statistics
multiple R 0.969842
R- square 0.94093
Normalized R-squared 0.906605
standard error 215.654
Observations 21

<table>
<thead>
<tr>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Relevance</th>
</tr>
</thead>
<tbody>
<tr>
<td>7</td>
<td>5,972,471</td>
<td>1,367,496</td>
<td>294,043</td>
<td>5,48E-07</td>
</tr>
<tr>
<td>13</td>
<td>6,045,865</td>
<td>465,566</td>
<td>215,654</td>
<td>1,017,706</td>
</tr>
<tr>
<td>20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3: Output multiple regression.

Table 4: Output and outcome.
3. Since -2.74615 ≤ -2.160368652, а 0.01666 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;

4. Since 3.84191 ≥ 2.160368652, а 0.002039 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;

5. Since -3.06239 ≤ -2.160368652, а 0.009081 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;

6. Since 5.164039 ≥ 2.160368652, а 0.000182 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;

7. Since 1.978199 ≤ 2.160368652, а 0.069497 ≥ 0.05, with 95% the reliability of the null hypothesis is rejected;

8. Since 3.233035 ≥ 2.160368652, а 0.006538 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected.

Limit values of F and t-statistics

1. Since 29.4043 ≥ 2.542144362 and 5.48E - 07 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

2. Since 7.458777 ≥ 1.978175101, а 4.77E - 06 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

3. Since -2.74615 ≤ -1.978175101, а 0.01666 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

4. Since 3.84191 ≥ 1.978175101, а 0.002039 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

5. Since -3.06239 ≤ -1.978175101, а 0.009081 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

6. Since 5.164039 ≥ 1.978175101, а 0.000182 ≤ 0.069497, with 93.0503% the reliability of the null hypothesis is rejected;

7. Since 1.978199 ≥ 1.978175101, а 0.069497 ≥ 0.069497, null hypothesis 93.0503% reliability NOW rejected;

8. As 3.233035 ≥ 2.160368652, а 0.006538 ≤ 0.0695, with 93.05% the reliability of the null hypothesis is rejected (Figures 9 and 10) (Tables 8-12).
Table 9: Analysis of variance.

<table>
<thead>
<tr>
<th></th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Relevance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>regression</td>
<td>7</td>
<td>9,572,471</td>
<td>1,367,496</td>
<td>294,043</td>
<td>5.48E-07</td>
</tr>
<tr>
<td>The residue</td>
<td>13</td>
<td>6,045,865</td>
<td>4,650,666</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>1,017,706</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 10: Analysis of variance.

<table>
<thead>
<tr>
<th></th>
<th>Odds</th>
<th>Standard errors</th>
<th>t- statistics</th>
<th>P- Value</th>
<th>Lower 95%</th>
<th>Top 95%</th>
<th>Lower 93.05%</th>
<th>Top 93.05%</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y- intersection</td>
<td>4,553,899</td>
<td>263,699</td>
<td>1,726,931</td>
<td>2.4E-10</td>
<td>3,984,212</td>
<td>5,123,586</td>
<td>4,032,257</td>
<td>5,075,542</td>
</tr>
<tr>
<td>The growth rate of GNP in the election year</td>
<td>0,700554</td>
<td>0,093923</td>
<td>7,458,777</td>
<td>4.7E-06</td>
<td>0,497645</td>
<td>0,903463</td>
<td>0,514757</td>
<td>0,886351</td>
</tr>
<tr>
<td>Abs. the value of the rate of inflation in an election year</td>
<td>-0,71228</td>
<td>0,259375</td>
<td>-274,615</td>
<td>0,0166</td>
<td>-0,396044</td>
<td>1,413,681</td>
<td>-0,19919</td>
<td>0,137,077</td>
</tr>
<tr>
<td>The number of quarters with GNP growth of &gt; 3.2%</td>
<td>0,904863</td>
<td>0,235524</td>
<td>384,191</td>
<td>0,00239</td>
<td>-0,396044</td>
<td>1,413,681</td>
<td>-0,19919</td>
<td>0,137,077</td>
</tr>
<tr>
<td>The number of periods spent ruling the party in power</td>
<td>-333,461</td>
<td>1,088,893</td>
<td>-306,239</td>
<td>0,009081</td>
<td>-568,703</td>
<td>-0,9822</td>
<td>-548,864</td>
<td>-118,059</td>
</tr>
<tr>
<td>Ruling party1 = republics - end</td>
<td>5,662,906</td>
<td>1,096,604</td>
<td>5,164,039</td>
<td>0,000182</td>
<td>3,293,837</td>
<td>8,031,975</td>
<td>3,493,631</td>
<td>7,832,181</td>
</tr>
<tr>
<td>War</td>
<td>4,714,529</td>
<td>2,383,243</td>
<td>1,978,199</td>
<td>0,069497</td>
<td>-0,43415</td>
<td>9,863,212</td>
<td>5.75e-05</td>
<td>9,429</td>
</tr>
<tr>
<td>Pull out whether the president for the next term?</td>
<td>3,983,671</td>
<td>1,232,177</td>
<td>3,223,035</td>
<td>0,006538</td>
<td>1,321,714</td>
<td>6,645,628</td>
<td>1,546,209</td>
<td>6,421,134</td>
</tr>
</tbody>
</table>

Table 11: Withdrawal of balance.

Table 12: Limit values of F and t- statistics.

<table>
<thead>
<tr>
<th>Regression statistics</th>
<th>df</th>
<th>SS</th>
<th>MS</th>
<th>F</th>
<th>Relevance F</th>
</tr>
</thead>
<tbody>
<tr>
<td>Multiple R</td>
<td>0,960578</td>
<td>0,22711</td>
<td>0,885986</td>
<td>2,370,322</td>
<td></td>
</tr>
</tbody>
</table>

Table 13: Output outcome.

Again repeat 29.4043 ≥ 2.542169216 and 5.48E-07 ≤ 0.069498, with 93.0503% the reliability of the null hypothesis is rejected. And further, since 1.978 ≥ 1.978, a 0,069497 ≤ 0,069497, null hypothesis 93.0503% reliability definitively rejected. Why is this important? Eventually chocolate weighs 100 grams, 95 or about 93? Consider the game of Russian roulette with one bullet in the drum. Do not we would like to see the drum, for example, 1,000 or a million rounds of ammunition? Where is the limit in medicine temperature 37°C? Almost all authors agree on the fact that 95% - a statistical threshold beyond which - a disease. And, secondly, if the counting carried out with the reverse end, the error is increased from 5% to about 7% and 15% (three times). That is, at 95% reliability a failure on 20 attempts, and at 93.0503% reliability a failure already at 14.38911032. When reliability of 85%, which is offered by Professor Winston, single failure accounted for 6.67 attempts. In a game of Russian roulette we get 20 and 14 and the 6-7 Chargers revolvers charged 1 cartridge. Even in the first case, we cannot risk it. Sachs [4] points out: "In special cases, especially when the test poses a risk to human life, you should take less than α = 0,001 probability of errors." In the second case, the risk has increased. In the third your chances are dangerous. The offender Gabbar Singh has played in Russian roulette, spinning drum six-shooter revolver. Thus, Mr. Gabbar would fulfill the conditions of a statistical reliability of 85%, if he had a 6.67 charging revolver with one cartridge (Figure 11).

We have no right to risk, said Holmes. Ten to one that they go down to the river, rather than up. And yet, we must take into account...
this possibility. Lips hero Sir Arthur waives 90% reliability.

In Tables 3-7 we have made the correction, whose outcome in Table 12. The last two columns are not the same way. In this way Table 3 contains, first, a mathematical error instead of the 95% reliability should indicate the reliability of 93.0502%. Second – under certain conditions and statistical. What conditions? If you are willing to risk – the coefficient of reliability of 93.0502% and 85% is statistically true. But if the risk is not acceptable – the coefficient of 4.714529 must be zero. Thus Tables 3 and 6 after the both patches reconstructed in Table 8. In support of Sax in [4] causes test the significance of the regression coefficient: "Checks null hypothesis H0: β ̂ = 0, i.e checks whether significantly different estimates of regression coefficients from zero. The border is set on the basis of the significance of the t-distribution [8].

\[
\begin{align*}
\hat{t} &= \frac{\hat{b}}{s_{\hat{b}}} \\
&= \frac{4.714529}{2.383243} \leq 1.978199034 \leq 2.160368652
\end{align*}
\]

C (n - 2) degrees of freedom. If statistics are greater than or equal to the limit of the significance of it, то β ̂ significantly different from zero".

Table 3 received:

<table>
<thead>
<tr>
<th>aWar</th>
<th>t</th>
<th>S</th>
<th>90%</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.714529</td>
<td>2.383243</td>
<td>21</td>
<td>95%</td>
</tr>
</tbody>
</table>

Coefficient aWar not significantly different from zero (P ≥ 0.05).

\[
y = 0.700554 \times x_1 - 0.71228 \times x_2 + 0.904863 \times x_3 - 3.33461 \times x_4 + 5.662906 \times x_5 + 4.714529 \times x_6 + 3.983671 \times x_7 + 45.53899
\]

(6)

In addition, Sachs [4] points out: Let's device consists of 300 complex elements. If these elements, for example, 284 completely smoothly, 12 have the reliability of 99%, and 4 – 98% reliability, the reliability of the device, provided independence is reliable elements.

\[
1,00284 \times 0,9912 \times 0,984 \times 1 \times 0,8864 \times 0,9224 = 0,8176 \approx 82%;
\]

This instrument no one would buy it".

Instead of equation 2, we can adopt a new mathematical model of the 6:

\[
\text{Projected\% of votes in the presidential election} = 48.60078 + 0.653815 \times \text{quarter growth GNP} - 0.97002 \times \text{abs. inf.} + 0.614881 \times \text{terms of power} + 5.028129 \times \text{Republican} + 4.167085 \times \text{President for a new term}
\]

(7)

After calculation not 21a and 22 members of the general population, you can get a little more accurate formula, the summary table to which you can build yourself.

\[
\text{Projected\% of votes in the presidential election} = 47.71879 + 0.661574 \times \text{quarter growth GNP} - 0.94685 \times \text{abs. inf.} + 0.621092 \times \text{quarter growth GNP} - 2.80492 \times \text{terms of power} + 5.214158 \times \text{Republican} + 4.669377 \times \text{President for a new term}
\]

(8)

Finally, Professor Winston in his work is not conducted analysis of residues (Figures 12-19) (Tables 12-17).

**Limit values of F and t- statistics**

1. Since 27.85619 ≥ 22.84772596 and 5.16E-07 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;
2. Since 6.543745 ≥ 2.144786681, a 1.3E-05 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;
3. Since -3.93503 ≤ - 2.144786681, a 0.001495 ≤ 0.05, with 95% the reliability of the null hypothesis is rejected;
4. Since 3.034645 ≥ 2.144786681, a 0.008917 ≤ 0.05, with 95% the
The ruling party 1 = Republicans

The schedule balances

Figure 16: Residual plot.

The growth rate of GNP in the election year

The schedule balances

Figure 17: Residual plot.

Pull out whether the President for the next term?

The schedule balances

Figure 18: Residual plot.

Abs. the value of the rate of inflation in an election year

The schedule balances

Figure 20: Residual plot.

Abs. the value of the rate of inflation in an election year

The schedule balances

Figure 21: Residual plot.

<table>
<thead>
<tr>
<th>Coefficient</th>
<th>standard error</th>
<th>t-Statistic</th>
<th>P-Value</th>
<th>Lower 95%</th>
<th>Top 95%</th>
<th>Lower 95.0%</th>
<th>Top 95.0%</th>
</tr>
</thead>
<tbody>
<tr>
<td>The growth rate of GNP in the election year</td>
<td>0.653815</td>
<td>0.099915</td>
<td>6,543.745</td>
<td>1.3E-05</td>
<td>0.43952</td>
<td>0.868111</td>
<td>0.43952</td>
</tr>
<tr>
<td>Abs. the value of the rate of inflation in an election year</td>
<td>-0.97002</td>
<td>0.24651</td>
<td>-393.503</td>
<td>0.001495</td>
<td>-149.873</td>
<td>-0.44131</td>
<td>-149.873</td>
</tr>
<tr>
<td>The number of quarters with GNP growth of &gt; 3.2%</td>
<td>0.614881</td>
<td>0.20262</td>
<td>3,034.645</td>
<td>0.008917</td>
<td>1,049.459</td>
<td>0.180303</td>
<td>1,049.459</td>
</tr>
<tr>
<td>The number of time spent by the ruling party in power</td>
<td>-320.944</td>
<td>1,194.815</td>
<td>-268.614</td>
<td>0.017731</td>
<td>-577.206</td>
<td>-0.64682</td>
<td>-577.206</td>
</tr>
<tr>
<td>Ruling party 1 = Republicans</td>
<td>5,028.129</td>
<td>1,152.555</td>
<td>4,362.593</td>
<td>0.00065</td>
<td>2,556.144</td>
<td>7,500,115</td>
<td>2,556,144</td>
</tr>
<tr>
<td>Pull out whether the president for the next term?</td>
<td>4,167.085</td>
<td>1,350.486</td>
<td>308.562</td>
<td>0.008059</td>
<td>1,270.582</td>
<td>7,063.589</td>
<td>1,270.582</td>
</tr>
</tbody>
</table>

Table 14: Analysis of variance.

Table 15: Analysis of variance.
reliability of the null hypothesis is rejected;

5. Since $-2.68614 \leq 2.144786681$, a 0.017731 $\leq 0.05$, with 95% the reliability of the null hypothesis is rejected;

6. Since 4.362593 $\geq 2.144786681$, a 0.00065 $\leq 0.05$, with 95% the reliability of the null hypothesis is rejected;

7. Since 3.08562 $\geq 2.144786681$, a 0.008059 $\leq 0.05$, with 95% the reliability of the null hypothesis is rejected.

More precisely will – to 98.2269% reliability (Figures 20-31).

Thus, having considered the multivariate regression model of
Figure 27: Standard normal probability calculator first depend ability, selected professor winston.

Figure 28: Standard normal probability calculator reliability, rejected holmes.

Figure 29: The standard normal probability calculator second liability, selected professor winston.

Figure 30: The standard normal probability calculator third threshold changes adopted by some researchers.
reliability must be in Table 3 instead of 95%;

2. 93.0503% reliability for this kind of data is sufficient and the model type 2 93.0503% reliability acceptable in life;

3. 93.0503% reliability, though not enough, but as there are managers as a risk-averse, and are not inclined, the first model to be adopted. Finally, we can recognize that the reliability of 93.0503% is inadequate and should be replaced by at least 95% reliability, and the model of the form (2) should lose factor $4\alpha$, and be re-constructed from the data in Table 7. The graphical analysis of residues also inclined to do so. The charts residues, relating to the Model (2) Professor Winston, the four of them has homoscedasticity. In the graphs relating to the model (6), we have constructed, homoscedasticity reduced to two, and the scatter of points is more favorable. Model (6) (as well as (7)) is 98.2269% reliability. Comparison of the standard normal probability plots gives the perception that a small reduction in reliability leads to a sharp change in the values of $z$-statistics and the number of standard deviations.

References

7. The Number empire.