The wild goose chase

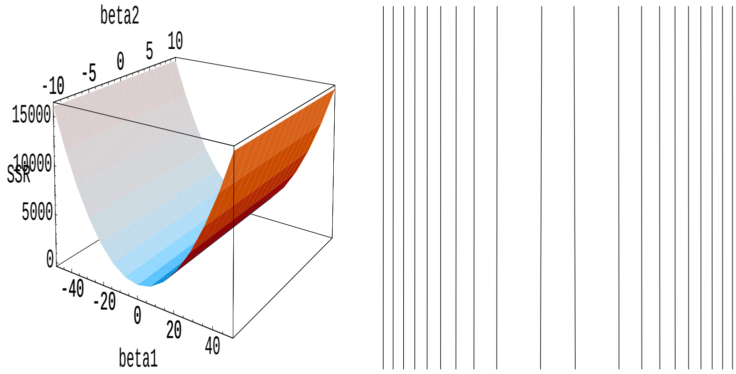
At some moment time . . . in 2004

Hope this few pages help you to figure out somehow what this ”staggering and gripping” stuff is about and, hopefully, to avoid fussing at Montse when it’s too late. Suggestions & comments are welcome at marti.mestieri@campus.uab.es. By the way, I haven’t managed to lighten up the subject (I tried it, though).

Preface

The aim of doing this tedious work is to fill the gap between the notes of the first semester and the notes on GLS Montse hanged on her webpage.
I’d like to think this notes are going to be a blessing in disguise for passing your exam.

Contents



Graphic representation of PM

1 Multicollinearity

Multicollinearity is the existence of linear relationship amongst the regressors. One can distinct between two different cases:

Perfect or Exact Multicollinearity: Regressors are linearly dependant. This is, given a sample space of regressors $\vec{x_1}, \ldots, \vec{x_k}$,

$$\exists \lambda_i \in \mathbb{R} \quad \forall i = 1, \ldots, k \mid \sum_{i=1}^{k} \lambda_i \vec{x_i} = 0 \quad \text{where } \exists \lambda_i \neq 0$$

Near Multicollinearity: This is a more general case. Given a sample space of regressors $\vec{x_1}, \ldots, \vec{x_k}$ and $\vec{v} \in \mathbb{R}^n$

$$\exists \lambda_i \in \mathbb{R} \quad \forall i = 1, \ldots, k \mid \sum_{i=1}^{k} \lambda_i \vec{x_i} = \vec{v}$$

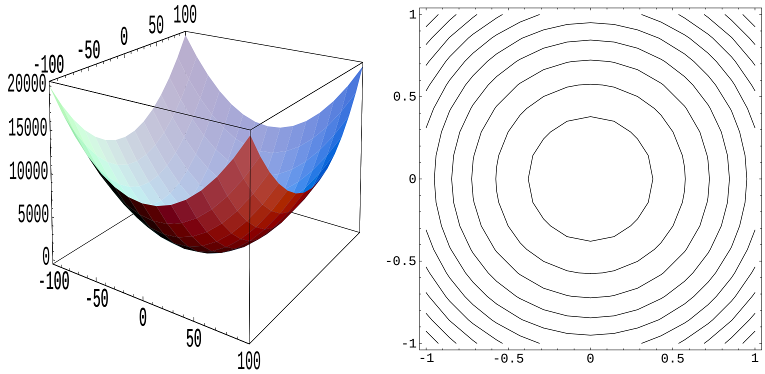
If $E(\vec{v}) = 0 \wedge V(\vec{v}) \equiv \sigma^2 = 0$, then NM turns out to be PM.

It is worth noting that $\vec{x_i}$ are (nx1) vectors and that $\vec{v}$ is interpreted as a (nx1) random vector.

1.1 Consequences of Multicollinearity

If we are dealing with perfect multicollinearity (PM), $(X'X)^{-1}$ is not defined. This is to say that there is not any privileged point to minimize the sum of squares.

On the other hand, if there is not multicollinearity, the shape of the SSR would be likely as shown.



Graphic representation of no multicollinearity

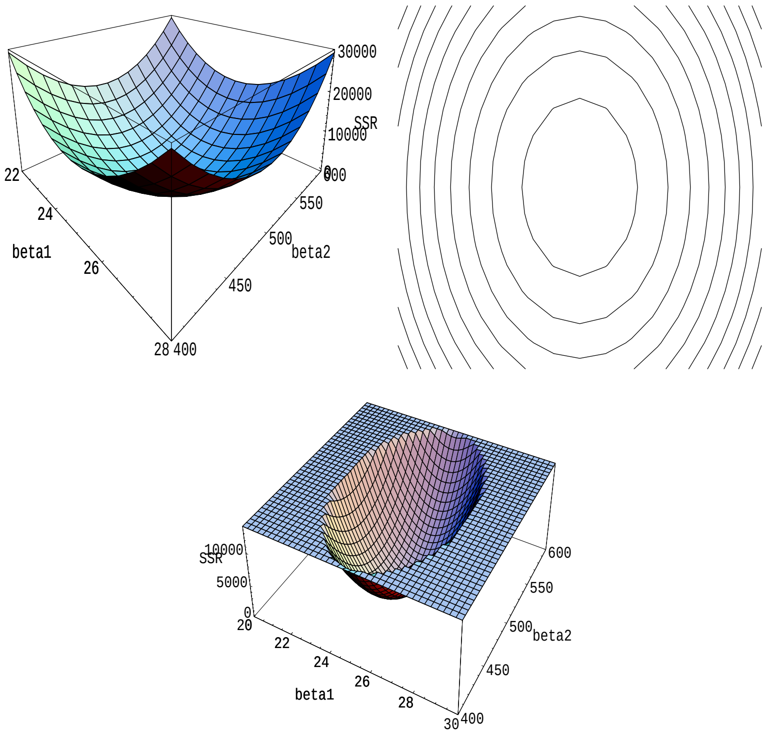
Things become harsher when near multicollinearity (NM) arises. Should this happen, the shape of the SSR would be an ellipsoid. In this case, $(X'X)^{-1}$ is large, thereby, so is $\widehat{V(\hat{\beta})}$. (Recall:$\widehat{V(\hat{\beta})} = \hat{\sigma}^2(X'X)^{-1}$). Taking this result into account implies one expects t-statistics closer to zero (ie, not significant). This low power of the significance hypothesis testing may lead easily to type II errors (not rejecting a false hypothesis). It is straightforward NM problem distorts (and somehow disrupts) the coefficients' interpretation. Nevertheless, R^2 and global significance tests *may be* very good.

1.2 Detection of NM

The list below points out a few outcomes which should warn us of NM

1. $det(X'X)$ is "close" to zero
2. High correlation amongst pairs of regressors is a sufficient condition for NM (pairwise relationship).
3. Coefficient instability: by dropping some observations, estimated coefficients do vary much.
4. High goodness of fit (R^2), but insignificant coefficients (low t-tests)
5. A mighty tool in order to detect NM is based upon its definition: the auxiliary regressions, this is, regress one regressor on the other ones. Mathematically, things go this way:

$$x_i = \frac{\lambda_1}{\lambda_i}x_1 + \ldots + \frac{\lambda_{i-1}}{\lambda_i}x_{i-1} + \frac{\lambda_{i+1}}{\lambda_i}x_{i+1} + \ldots + \frac{\lambda_k}{\lambda_i}x_k + \frac{v}{\lambda_i}$$



Graphic representation of NM

This regressions, sometimes, allow us to find out whether exists or not NM. For instance, if the regression is globally significant.

1.3 Corrective Measures

If we are only interested in prediction, regarding the high R^2, we should consider doing nothing as our best choice. Nevertheless, errors may arise. Alternatively, some recipes exist to lower the effects of NM.

***A priori* information**: setting extra restrictions on the model based upon previous experience or data and common sense.

Mixing cross series data and time series data:one can consider time series observations to predict current behaviours and, so restrict the cross section model.

Ruling out some variables: If we drop one variable, even though the improvement in the coefficients' significance, there is no way to maintain the estimation unbiased and specification errors arise. Further details in the next section.

Transformation of the variables:It does exist several transformations in order to reduce NM. Arguably, the most common are the first difference transformation for time series and the "rate transformation" (check it out).

Adding new or extra data: As long as NM is a problem of the data set, modifying it should improve the set's quality. Spreading the width of our data set is sometimes enough. If not, we shall consider obtaining a new one.

2 Specification Errors

As we introduced it above, specification errors come up whether relevant variables are omitted or irrelevant variables are included in the econometric model. I am going to deal with these two problems separately. Firstly, I will focus on omitting relevant variables.

2.1 Omission of relevant variables

To begin with, in this subsection it is assumed a true classical model such that:

$$Y = \underbrace{X}_{nxk} \cdot \underbrace{\beta}_{kx1} + \underbrace{\omega}_{nxs} \cdot \underbrace{\gamma}_{sx1} + \varepsilon$$

However, a short-sighted econometrician hasn't heeded the reality and sets the model to estimate as follows:

$$Y = X \cdot \beta + u \quad \text{where } u \text{ stands for the error term}$$

It is assumed all the classical assumptions are satisfied.
One can estimate the true model via least squares (as we have been doing so far), thus,

$$\hat{\beta}_{OLS} = (X'X)^{-1}(X'Y) = (X'X)^{-1}(X'(X\beta + \omega\gamma + \varepsilon))$$

$$\hat{\beta}_{OLS} = \beta + (X'X)^{-1}X'\omega\gamma + (X'X)^{-1}X'\varepsilon$$

Hence, the expectation of $\hat{\beta}_{OLS}$ is:

$$E(\hat{\beta}_{OLS}) = \beta + E((X'X)^{-1}X'\omega\gamma) + 0 \quad \text{since } E(\varepsilon) = 0$$

Taking this result into account, it is immediate that the econometrician is going to obtain a biased estimation of β unless $E((X'X)^{-1}X'\omega\gamma)$, which is nearly impossible (this would happen if X's and ω were independent, this is, no correlated at all).
Similarly, this happens as well to the variance, the econometrician is going to estimate a biased $\hat{\sigma}^2$. Actually,

$$E(\hat{\sigma}^2) = \sigma^2 + A$$

And A being positive or negative depends on the correlation between X's an ω's
It is worth noting that in this cases neither the $\beta's$ nor the σ^2 will be distributed as usually (t-statistic, etcetera). It is true even if X's and ω are independent, because hypothesis tests are not true.

2.2 Inclusion of irrelevant variables

On the contrary, the econometrician could set a model

$$Y = X\beta + \Omega\gamma + u \qquad u \sim (\vec{0}, \sigma^2 Id)$$

when the true model is ($\gamma = 0$)

$$Y = X\beta + \varepsilon \qquad \varepsilon \sim (\vec{0}, \zeta^2 Id)$$

This is to say, irrelevant variables have been included in the econometric model.
Defining $Z \equiv (X|\ \Omega)$ and $\theta \equiv \begin{pmatrix} \beta \\ \gamma \end{pmatrix}$ the econometric model to estimate becomes:

$$Y = Z\theta + u$$

Hence, applying the OLS estimation

$$\hat{\theta} = \begin{pmatrix} \hat{\beta} \\ \hat{\gamma} \end{pmatrix} = (Z'Z)^{-1}Z'Y$$

$$E(\hat{\theta}) = \begin{pmatrix} E(\hat{\beta}) \\ E(\hat{\gamma}) \end{pmatrix} = \begin{pmatrix} \beta \\ \gamma \end{pmatrix} = \begin{pmatrix} \beta \\ 0 \end{pmatrix}$$

So, $\hat{\theta}$ is unbiased. Furthermore, the estimation of variance is unbiased:

$$E(\hat{\sigma}^2) = \sigma^2$$

Nevertheless, it can be proven that variance of $\hat{\theta}$ is not efficient

$$Var(\hat{\theta}) \geq Var(\hat{\beta})$$

Summing up,

Omission of relevant variables	Inclusion of irrelevant variables
$\hat{\beta}$ biased $\hat{\sigma}^2$ biased Tests not invalid	$\hat{\beta}$ unbiased but inefficient $\hat{\sigma}^2$ unbiased Tests valid (but less power)