Module 2 – Regression and Prediction

CASE STUDY ACTIVITY TUTORIAL

Case Study 3: Do poor countries grow faster than rich countries?

MIT PROFESSIONAL EDUCATION
Digital Programs

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

Victor Chernozhukov (MIT)

June 4, 2016

2016-06-29

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

Victor Chernozhukov (MIT)

June 4, 2016

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

Introduction

- We provide an empirical example of partialling-out with Lasso to estimate the regression coefficient β_1 in the high-dimensional linear regression model:

$$Y = \beta_1 D + \beta_2' W + \epsilon.$$

- Specifically we are interested in how the rates at which economies of different countries grow (Y) are related to the initial wealth levels in each country (D) controlling for country's institutional, educational, and other similar characteristics (W).

- The relationship is captured by β_1, the "speed of convergence/divergence", which measures the speed at which poor countries catch up ($\beta_1 < 0$) or fall behind ($\beta_1 > 0$) rich countries, after controlling for W.

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

└─Introduction

- We provide an empirical example of partialling-out with Lasso to estimate the regression coefficient β_1 in the high-dimensional linear regression model:

$$Y = \beta_1 D + \beta_2' W + \epsilon.$$

- Specifically we are interested in how the rates at which economies of different countries grow (Y) are related to the initial wealth levels in each country (D) controlling for country's institutional, educational, and other similar characteristics (W).
- The relationship is captured by β_1, the "speed of convergence/divergence", which measures the speed at which poor countries catch up ($\beta_1 < 0$) or fall behind ($\beta_1 > 0$) rich countries, after controlling for W.

In this segment, we provide an empirical example of using partialling-out with Lasso to estimate the regression coefficient β_1 in the high-dimensional linear regression model:

$$Y = \beta_1 D + \beta_2 W + \epsilon.$$

Specifically we are interested in how the rates at which economies of different countries grow, denoted by Y, are related to the initial wealth levels in each country, denoted by D, controlling for country's institutional, educational, and other similar characteristics, denoted by W.

The relationship is captured by the regression coefficient β_1.

In this example, this coefficient is called the "speed of convergence/divergence", as it measures the speed at which poor countries catch up or fall behind wealthy countries, controlling for W.

- Our inference question here is: do poor countries grow faster than rich countries, controlling for educational and other characteristics? In other words, is the speed of convergence negative:

$$-_1 < 0?$$

- This is the Convergence Hypothesis predicted by the Solow Growth Model

- Robert M. Solow is a world-renowned MIT economist who won the Nobel Prize in Economics.

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

- Our inference question here is: do poor countries grow faster than rich countries, controlling for educational and other characteristics? In other words, is the speed of convergence negative:

$$-_1 < 0?$$

- This is the Convergence Hypothesis predicted by the Solow Growth Model
- Robert M. Solow is a world-renowned MIT economist who won the Nobel Prize in Economics.

So, our inference question here is: do poor countries growth faster than rich countries, controlling for educational and institutional characteristics? In other words, is the speed of convergence negative, namely, is $-_1$ negative?

In Economics this question is known as the Convergence Hypothesis, which is predicted by the Solow Growth Model.

The growth model was developed by Professor Robert M. Solow, a world-renowned MIT economist who won the Nobel Prize in Economics in 1987.

Barro-Lee Growth Data

- The outcome (Y) is the realized annual growth rate of a country's wealth (Gross Domestic Product per capita).
- The target regressor (D) is the initial level of the country's wealth
- The target parameter — is the speed of convergence, which measures the speed at which poor countries catch up with rich countries.
- The controls (W) include measures of education levels, quality of institutions, trade openness, and political stability in the country.

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

Barro-Lee Growth Data

Barro-Lee Growth Data

- The outcome (Y) is the realized annual growth rate of a country's wealth (Gross Domestic Product per capita).
- The target regressor (D) is the initial level of the country's wealth
- The target parameter — is the speed of convergence, which measures the speed at which poor countries catch up with rich countries.
- The controls (W) include measures of education levels, quality of institutions, trade openness, and political stability in the country.

In this case study, we use the data collected by economists Robert Barro and Jong-Wha Lee.

In this data set, the outcome Y is the realized annual growth rate of a country's wealth, measured by the Gross Domestic Product per capita.

The target regressor (D) is the initial level of the country's wealth

The controls (W) include measures of education levels, quality of institutions, trade openness, and political stability in the country.

High-Dimensional Setting

- The sample contains 90 countries and about 60 controls. Thus

$$p \approx 60, \quad n = 90$$

  and p/n is not small.

- We expect the least squares method to provide a poor/noisy estimate of β_1.

- We expect the method based on partialling-out with Lasso to provide a high quality estimate of β_1.

- The sample contains 90 countries and about 60 controls. Thus

  p ¥ 60, n = 90

  and p/n is not small.
- We expect the least squares method to provide a poor/noisy estimate of —$_1$.
- We expect the method based on partialling-out with Lasso to provide a high quality estimate of —$_1$.

The data set contains 90 countries and about 60 controls. Thus p is approximately 60 and n = 90 and p over n is not small. This means that we operate in the high-dimensional setting.

Therefore, we expect the least squares method to provide a poor, very noisy estimate of —$_1$.

In contrast, we expect the method based on partialling-out with Lasso to provide a high quality estimate of —$_1$.

Results

	Estimate	Std. Error	95% Conf. Interval
Least squares	-0.009	0.030	[-0.069, 0.050]
Partialling-out via lasso	-0.044	0.015	[-0.075, -0.014]

- As expected, least squares provides a rather noisy estimate of the speed of convergence, and does not allow us to answer the question about the convergence hypothesis.
- In sharp contrast, partialling-out via Lasso provides a more precise estimate.
- The lasso based point estimate is ≠4% and the 95% confidence interval for the (annual) rate of convergence is ≠7.5% to ≠1.5%.
- This empirical evidence does support the convergence hypothesis.

Regression 2.4. Case Study: Do poor countries grow faster than rich countries?

└─Results

Results

	Estimate	Std. Error	95% Conf. Interval
Least squares	-0.009	0.030	[-0.069, 0.050]
Partialling-out via lasso	-0.044	0.015	[-0.075, -0.014]

- As expected, least squares provides a rather noisy estimate of the speed of convergence, and does not allow us to answer the question about the convergence hypothesis.
- In sharp contrast, partialling-out via Lasso provides a more precise estimate.
- The lasso based point estimate is ≠4% and the 95% confidence interval for the (annual) rate of convergence is ≠7.5% to ≠1.5%.
- This empirical evidence does support the convergence hypothesis.

We now present the empirical results in a table that you can see here.

The table shows the estimates of the speed of convergence —ı obtained by least squares, and by partialling out with lasso. The table also provides the standard error and 95% confidence interval.

As we expected, the least squares method provides a rather noisy estimate of the annual speed of convergence, and does not allow us to answer the question about the convergence hypothesis.

In sharp contrast, partialling-out via Lasso provides a much more precise estimate and does support the convergence hypothesis. We see that the lasso based estimate of —ı is ≠4% and the 95% confidence interval for the (annual) rate of convergence is from ≠7.5% to ≠1.5%. Now... this empirical evidence does support the convergence hypothesis.

Summary

- In this segment, we have examined an empirical example in the high-dimensional setting.
- Least squares yields a very noisy estimate of the target regression coe‰cient and does not allow us to answer an important empirical question
- Lasso does yield a precise estimate of the regression coe‰cient and does allow us to answer that question
- We have found significant empirical evidence supporting the convergence hypothesis of Solow.

Summary

- In this segment, we have examined an empirical example in the high-dimensional setting.
- Least squares yields a very noisy estimate of the target regression coe‰cient and does not allow us to answer an important empirical question
- Lasso does yield a precise estimate of the regression coe‰cient and does allow us to answer that question
- We have found significant empirical evidence supporting the convergence hypothesis of Solow.

Let us summarize.

In this segment, we have examined an empirical example in the high-dimensional setting.

Using least squares in this setting gives us a very noisy estimate of the target regression coe‰cient and does not allow us to answer an important empirical question.

In sharp contrast, using the partialling-out method with Lasso does give us a precise estimate of the regression coe‰cient and does allow us to answer that question.

We have found significant empirical evidence supporting the convergence hypothesis of Solow.

Module 2 – Regression and Prediction

THANK YOU

Case Study 3: Do poor countries grow faster than rich countries?

PROFESSIONAL EDUCATION
Digital Programs