Introduction

Definition of the constrained linear multiple regression model

The constrained ordinary least squares estimator
- Estimation of $\beta_0$
- Estimation of $\sigma_0^2$

Constrained maximum likelihood estimator

Statistical properties
- Unbiasedness
- Variance-covariance matrix of the $\hat{\beta}_{COLS}$
- Efficiency
- Large sample properties

Summary

Appendix
Consider the multiple linear regression model:

\[ Y = X\beta_0 + u \]

subject to a set of linear constraints

\[ R\beta_0 = q \]

where

- \( Y \in \mathbb{R}^n \), \( X \) is a matrix of fixed and random variables and is of full column rank, \( \beta_0 \in \mathbb{R}^k \), \( \mathbb{E}(u \mid X) = 0_{n \times 1} \), and \( V(u \mid X) = \sigma_0^2 I_n \);

- \( R \in \mathcal{M}_{p \times k} \), \( \text{Rk}(R) = p \), and \( q \in \mathcal{M}_{p \times 1} \).
The constrained model can be estimated with the least squares method or the maximum likelihood principle:

- The constrained ordinary least squares estimator (COLS);
- The constrained maximum likelihood estimator (CML).

Main objectives of Part I:

- How can we define the constrained ordinary least squares estimator and the constrained maximum likelihood estimator?
- What are their main statistical properties?
2. Definition of the constrained linear multiple regression model

**Definition**

Let $Y \in \mathbb{R}^n$ denote the dependent variable, $X \in M_{n,k}$ denote a matrix of constant and random regressors, $R \in M_{p,k}$ denote a fixed matrix and $q \in M_{p,1}$ denote a fixed vector. The constrained linear multiple regression model is defined to be:

$$Y = X\beta_0 + u$$

s.t. $R\beta_0 = q$

with $\mathbb{E}(u \mid X) = 0_{n \times 1}$, $\mathbb{V}(u \mid X) = \sigma_0^2 I_n$, $Rk\left(\mathbb{E}X \left[ x_i x_i' \right]\right) = k$ or $Rk(X) = k$, $x_i'$ is the $i^{th}$ row of $X$, and $Rk(R) = p$ with $p \leq k$. 

□
3. The constrained ordinary least squares estimator

3.1. Estimation of $\beta_0$

**Definition**

Under suitable regularity conditions, the constrained ordinary least squares estimator of $\beta_0$, denoted $\hat{\beta}_{COLS}$, is the solution of the following constrained minimization program:

$$\min_{\beta_0} \| Y - X \beta_0 \|_2^2$$

subject to $R\beta_0 = q$.

**Proof:** See Appendix 1.
Proposition

Under suitable regularity conditions, the constrained ordinary least squares estimator of $\beta$ is given by:

$$\hat{\beta}_{COLS} = \hat{\beta}_{OLS} - (X'X)^{-1} R' [R(X'X)^{-1} R']^{-1} [R\hat{\beta}_{OLS} - q]$$
3.2. Estimator of $\sigma_0^2$

**Proposition**

The fitted vector of residuals obtained from the constrained ordinary least squares estimator can be decomposed as the sum of two orthogonal vectors:

$$\hat{u}_{COLS} = \hat{u}_{OLS} + \hat{v}$$

where

$$\hat{v} = X \left( \hat{\beta}_{OLS} - \hat{\beta}_{COLS} \right)$$

$$\begin{align*}
\hat{v} &= X (X'X)^{-1} R' \left[ R(X'X)^{-1} R' \right]^{-1} R(X'X)^{-1} X' u \\
&= \mathbb{P} X (X'X)^{-1} R' u
\end{align*}$$
Proposition

An unbiased estimator of $\sigma_0^2$ is defined to be:

$$\hat{\sigma}^2_{COLS} = \frac{||\hat{u}_{COLS}||^2_n}{n - k + p}$$

Proof: See Appendix 2.
4. Constrained maximum likelihood estimator

**Definition**
Under suitable regularity conditions, the constrained maximum likelihood estimator of \( \theta_0 = (\beta_0', \sigma_0^2)' \) is the solution of the following constrained maximization problem:

\[
\max_{\theta_0} \ell(y \mid x; \theta_0) \\
\text{s.t. } R\beta_0 = q.
\]
Definition

Let \( \mathcal{L} \) denote the Lagrangian associated to the constrained maximization problem:

\[
\mathcal{L} (\theta_0; \lambda) = \ell (y \mid x; \theta_0) - \lambda' (R\beta_0 - q).
\]

Under suitable regularity conditions, a solution (constrained maximum likelihood estimate) to this problem satisfies:

\[
\begin{aligned}
\frac{\partial \ell}{\partial \beta_0} (\cdot; \hat{\beta}_{CML}, \hat{\sigma}^2_{CML}) - R' \lambda &= 0_{k \times 1} \\
\frac{\partial \ell}{\partial \sigma^2_0} (\cdot; \hat{\beta}_{CML}, \hat{\sigma}^2_{CML}) &= 0 \\
R\hat{\beta}_{CML} - q &= 0_{p \times 1}.
\end{aligned}
\]
Proposition

Under suitable regularity conditions, the constrained maximum likelihood estimator of $\theta_0 = (\beta'_0, \sigma^2_0)'$, is defined to be:

$$\hat{\beta}_{CML} = \hat{\beta}_{COLS}$$

$$\hat{\sigma}^2_{CML} = \frac{1}{n} \left\| Y - X\hat{\beta}_{CML} \right\|^2.$$

Constrained maximum likelihood estimator
5. Statistical properties

5.1. Unbiasedness property

**Proposition**

If the model is correctly specified—\( \beta \) satisfies the set of linear constraints \( R\beta = q \)—and the exogeneity assumption holds, then the constrained ordinary least squares estimator of \( \beta \), \( \hat{\beta}_{COLS} \), is unbiased:

\[
E \left[ \hat{\beta}_{COLS} \right] = \beta \text{ for all } \beta \in \Theta.
\]

\[\Box\]

**Proof:** See Appendix 3.
Proposition

If the model is incorrectly specified—\( \beta \) does not satisfy the set of linear constraints \( R\beta = q \), the constrained ordinary least squares estimator of \( \beta \) is biased.

Proof: In this case,

\[
\mathbb{E} \left[ R\hat{\beta}_{OLS} - q \mid X \right] \neq 0_{p \times 1}
\]

and the result follows. □
5.2. Variance-covariance matrix of $\hat{\beta}_{COLS}$

Proposition

The variance-covariance matrix of the constrained ordinary least squares estimator of $\beta$ is defined to be:

$$\mathbb{V}
\left(\beta_{COLS} \mid X\right)
= \sigma^2 (X'X)^{-1} \left[I_k - R'[R(X'X)^{-1}R']^{-1}R(X'X)^{-1}\right]$$

and

$$\mathbb{V}
\left(\beta_{COLS} \mid X\right) \preceq \mathbb{V}
\left(\beta_{OLS} \mid X\right)$$
Implications...

1. The constrained ordinary least squares estimator might be biased, especially if the model is incorrectly specified.

2. The constraint ordinary least squares estimator is more precise than the ordinary least squares estimator:

   \[ \text{Var} \left( \hat{\beta}_{COLS} \mid X \right) \preceq \text{Var} \left( \hat{\beta}_{OLS} \mid X \right) \]

3. There is a bias-efficiency trade-off: introducing more constraints in the model improves the precision of the estimates but the corresponding estimates are more likely biased! The converse is also true.
5.3. Efficiency

Efficiency can be shown as in Chapter 2 with:

- The Gauss-Markov theorem (semi-parametric model);
- The FDCR lower bound (parametric model).
5.4. Large sample properties
(a) Parametric model

Proposition

Under suitable regularity conditions, the sampling distribution of the constrained ordinary least squares estimator of $\beta_0$ in a Gaussian constrained linear multiple regression model is:

$$\hat{\beta}_{COLS} | X \sim N \left( \beta_0, \sigma^2_0 (X'X)^{-1} \left[ I_k - R'[R(X'X)^{-1}R']^{-1} R(X'X)^{-1} \right] \right)$$
(b) Semi-parametric model

Proposition (Independent and identically distributed observations)

Under suitable regularity conditions, the sampling distribution of the constrained ordinary least squares estimator of $\beta_0$ in a semi-parametric constrained linear multiple regression model (with random regressors) is:

$$\hat{\beta}_{COLS} \overset{d}{\sim} N\left(\beta_0, \frac{\sigma_0^2}{n}\left[\mathbb{E}_X (x_i x_i')\right]^{-1}\left\{ I_k - R' \left[ R \left[\mathbb{E}_X (x_i x_i')\right]^{-1} R'\right]^{-1} R \left[\mathbb{E}_X (x_i x_i')\right]^{-1}\right\}\right)$$

with

$$y_i = x'_i \beta_0 + u_i$$

$$s.t. \ R\beta_0 = q.$$  

where $\mathbb{E} (u_i \mid X) = 0$, $V (u \mid X) = \sigma_0^2 I_n$, and $\mathbb{E}_X (x_i x_i')$ is nonsingular for all $i = 1, \ldots, n$. □
6. Summary

- What is the definition of the constrained multiple linear regression model? What assumptions does one need?

- Provide simple examples (from economic theory or finance theory) of constrained models.

- How can one derive the constrained ordinary least squares estimator? The constrained maximum likelihood estimator?

- What are the main statistical properties (in terms of unbiasedness and efficiency) of the constrained ordinary least squares estimator? Under which critical assumptions?
Appendix 1: Derivation of the constrained OLS estimator.

■ **Step 1**: Writing the Lagrangian...

\[ \mathcal{L}(\beta_0; \lambda) = \| Y - X\beta_0 \|^2 + \lambda'(R\beta_0 - q) \]

where \( \lambda \in \mathcal{M}_{p,1} \) is the vector of Lagrange multipliers.

■ **Step 2**: Derivation of the first-order conditions

\[
\begin{cases}
\frac{\partial \mathcal{L}}{\partial \beta_0}(\cdot) = 0_{k \times 1} \\
\frac{\partial \mathcal{L}}{\partial \lambda}(\cdot) = 0_{p \times 1}
\end{cases}
\iff
\begin{cases}
-2X' \left( Y - X\hat{\beta}_{COLS} \right) + R'\hat{\lambda} = 0_{k \times 1} \\
R\hat{\beta}_{COLS} - q = 0_{p \times 1}
\end{cases}
\]

(1)
Step 3: Use Eq. (1) to express $\hat{\beta}_{COLS}$ as a function of $\lambda$

$$\hat{\beta}_{COLS} = (X'X)^{-1}(X'Y) - \frac{1}{2}(X'X)^{-1}R'\hat{\lambda}$$

or

$$\hat{\beta}_{COLS} = \hat{\beta}_{OLS} - \frac{1}{2}(X'X)^{-1}R'\hat{\lambda} \quad (3)$$

Step 4: Find the expression of $\lambda$ after replacing the previous expression in Eq. (2)

$$R\hat{\beta}_{COLS} = q \iff R\hat{\beta}_{OLS} - \frac{1}{2}R(X'X)^{-1}R'\hat{\lambda} = q$$

i.e.,

$$\hat{\lambda} = 2[R(X'X)^{-1}R']^{-1}[R\hat{\beta}_{OLS} - q]$$

Step 5: Find the expression of $\hat{\beta}_{COLS}$ using Eq. (3)

$$\hat{\beta}_{COLS} = \hat{\beta}_{OLS} - (X'X)^{-1}R'[R(X'X)^{-1}R']^{-1}[R\hat{\beta}_{OLS} - q]$$
Appendix 2a: Decomposition of $\hat{u}_{COLS}$

- **Step 1**: Definition of $\hat{u}_{COLS}$
  \[ \hat{u}_{COLS} = Y - \hat{Y}_{COLS} = Y - X\beta_{COLS}. \]

- **Step 2**: Insert $Y_{OLS}!$
  \[ \hat{u}_{COLS} = Y - \hat{Y}_{OLS} + \hat{Y}_{OLS} - \hat{Y}_{COLS} \]
  \[ = \hat{u}_{OLS} + \hat{Y}_{OLS} - \hat{Y}_{COLS} = \hat{u}_{OLS} + \hat{v} \]

  with $\hat{v} = X\left(\hat{\beta}_{OLS} - \hat{\beta}_{COLS}\right)$ and $\hat{u}_{OLS} \perp \hat{v}$ (why?).
Step 3: Use the definition of the constrained ordinary least squares estimator to rewrite \( \hat{v} \):

\[
\hat{v} = \mathbf{X} \left\{ (\mathbf{X}'\mathbf{X})^{-1} R' \left[ R(\mathbf{X}'\mathbf{X})^{-1} R' \right]^{-1} (\mathbf{R}\hat{\mathbf{\beta}}_{OLS} - q) \right\}
\]

with:

\[
\mathbf{R}\hat{\mathbf{\beta}}_{OLS} - q = \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{Y} - q \\
= \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'(\mathbf{X}\mathbf{\beta} + \mathbf{u}) - q \\
= \mathbf{R}\mathbf{\beta}_0 - q + \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{u} = \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{u}.
\]

Step 4: Rearrange terms!

\[
\hat{v} = \mathbf{X} \left\{ (\mathbf{X}'\mathbf{X})^{-1} R' \left[ R(\mathbf{X}'\mathbf{X})^{-1} R' \right]^{-1} (\mathbf{R}\hat{\mathbf{\beta}}_{OLS} - q) \right\}
\]

\[
= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R' \left[ R(\mathbf{X}'\mathbf{X})^{-1} R' \right]^{-1} \mathbf{R}(\mathbf{X}'\mathbf{X})^{-1} \mathbf{X}'\mathbf{u}
\]

\[
= \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R' \left[ (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R')' (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R') \right]^{-1} (\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R')' \mathbf{u}
\]

\[
= \mathbf{P} \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1} R' \mathbf{U}
\]
Appendix 2b: Unbiased estimator of $\sigma_0^2$

- **Step 1**: Use the orthogonal decomposition of $\hat{u}_{COLS}$

$$\|\hat{u}_{COLS}\|_{I_n}^2 = \|\hat{u}_{OLS}\|_{I_n}^2 + \|\hat{v}\|_{I_n}^2.$$ 

- **Step 2**: Use the linearity of the expectation operator

$$E\left[\|\hat{u}_{COLS}\|_{I_n}^2 \mid X\right] = E\left[\|\hat{u}_{OLS}\|_{I_n}^2 \mid X\right] + E\left[\|\hat{v}\|_{I_n}^2 \mid X\right]$$

- **Step 3**: Use the two property (i) $\text{Tr}(a) = a$ for all $a \in \mathbb{R}$ and (ii) $E[\text{Tr}(.))] = \text{Tr}[E(.)]$

$$E\left[\|\hat{u}_{COLS}\|_{I_n}^2 \mid X\right] = \sigma_0^2 \left((n - k) + \text{Tr}(P_{XX^{-1}R^t})\right)$$

with:

$$\text{Tr}\left(P_{XX^{-1}R^t}\right) = p.$$ 

- **Step 4**: Conclude!

$$E\left[\|\hat{u}_{COLS}\|_{I_n}^2 \mid X\right] = \sigma_0^2(n - k + p).$$
Appendix 3: Proof of unbiasedness

Taking the conditional expectation with respect to $X$ of the constrained ordinary least squares estimator of $\beta$, one gets:

$$
E \left[ \hat{\beta}_{COLS} \mid X \right] = E \left[ \hat{\beta}_{OLS} \mid X \right] - (X'X)^{-1} R' R (X'X)^{-1} R' \left[ R \hat{\beta}_{OLS} - q \mid X \right].
$$

Therefore,

$$
E \left[ R \hat{\beta}_{OLS} - q \mid X \right] = R(X'X)^{-1} X' E [u \mid X] + R\beta - q = 0.
$$

The first right-hand side term equals zero due to the exogeneity assumption whereas the second term equals zero since the model is correctly specified and the null assumption holds.

Since $E \left[ \hat{\beta}_{OLS} \mid X \right] = \beta$, one has:

$$
E \left[ \hat{\beta}_{COLS} \mid X \right] = \beta_0. \square
$$