Remark: The solution of the above minimization problem is equivalent to the GLE estimator of β in the following regression model:

$$Z'y = Z'X\beta + Z'u,$$

where Z, y, X, β and u are $n \times r$, $n \times 1$, $n \times k$, $k \times 1$ and $n \times 1$ matrices or vectors. Note that $r > k$.

$y^* = Z'y$, $X^* = Z'X$ and $u^* = Z'u$ denote $r \times 1$, $r \times k$ and $r \times 1$ matrices or vectors, where $r > k$.

Rewrite as follows:

$$y^* = X^*\beta + u^*,$$

$\Longrightarrow$ r is taken as the sample size.

u^* is a $r \times 1$ vector.

The elements of u^* are correlated with each other, beacuse each element of u^* is a function of $u_1, u_2, \cdots, u_n$.

The variance of u^* is:

$$\mathrm{V}(u^*) = \mathrm{V}(Z'u) = \sigma^2 Z'\Omega Z.$$

Go back to GMM:

$$\begin{aligned}
&(y - X\beta)'Z(Z'\Omega Z)^{-1}Z'(y - X\beta) \\
&= y'Z(Z'\Omega Z)^{-1}Z'y - \beta'X'Z(Z'\Omega Z)^{-1}Z'y - y'Z(Z'\Omega Z)^{-1}Z'X\beta + \beta'X'Z(Z'\Omega Z)^{-1}Z'X\beta \\
&= y'ZWZ'y - 2y'Z(Z'\Omega Z)^{-1}Z'X\beta + \beta'X'Z(Z'\Omega Z)^{-1}Z'X\beta.
\end{aligned}$$

Note that $\beta'X'Z(Z'\Omega Z)^{-1}Z'y = y'Z(Z'\Omega Z)^{-1}Z'X\beta$ because both sides are scalars.

Remember that $\dfrac{\partial Ax}{x} = A'$ and $\dfrac{\partial x'Ax}{x} = (A + A')x$.

Then, we obtain the following derivation:

$$
\begin{aligned}
&\frac{\partial(y - X\beta)'Z(Z'\Omega Z)^{-1}Z'(y - X\beta)}{\partial\beta} \\
&\qquad = -2(y'Z(Z'\Omega Z)^{-1}Z'X)' + (X'Z(Z'\Omega Z)^{-1}Z'X + (X'Z(Z'\Omega Z)^{-1}Z'X)')\beta \\
&\qquad = -2X'Z(Z'\Omega Z)^{-1}Z'y + 2X'Z(Z'\Omega Z)^{-1}Z'X\beta = 0
\end{aligned}
$$

The solution of β is denoted by β_{GMM}, which is:

$$
\beta_{GMM} = (X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'y.
$$

The mean of β_{GMM} is asymptotically obtained.

$$
\begin{aligned}
\beta_{GMM} &= (X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'(X\beta + u) \\
&= \beta + (X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'u \\
&= \beta + \Big((\frac{1}{n}X'Z)(\frac{1}{n}Z'\Omega Z)^{-1}(\frac{1}{n}Z'X)\Big)^{-1}(\frac{1}{n}X'Z)(\frac{1}{n}Z'\Omega Z)^{-1}(\frac{1}{n}Z'u)
\end{aligned}
$$

We assume that

$$\frac{1}{n}X'Z \longrightarrow M_{xz} \quad \text{and} \quad \frac{1}{n}Z'\Omega Z \longrightarrow M_{z\Omega z},$$

which are $k \times r$ and $r \times r$ matrices.

From the assumption of $\frac{1}{n}Z'u \longrightarrow 0$, we have the following result:

$$\beta_{GMM} \longrightarrow \beta + (M_{xz}M_{z\Omega z}^{-1}M_{xz}')^{-1}M_{xz}M_{z\Omega z}^{-1} \times 0 = \beta.$$

Thus, β_{GMM} is a consistent estimator of β (i.e., asymptotically unbiased estimator).

The variance of β_{GMM} is asymptotically obtained as follows:

$$
\begin{aligned}
\mathrm{V}(\beta_{GMM}) &= \mathrm{E}\Big((\beta_{GMM} - \mathrm{E}(\beta_{GMM}))(\beta_{GMM} - \mathrm{E}(\beta_{GMM}))'\Big) \approx \mathrm{E}\Big((\beta_{GMM} - \beta)(\beta_{GMM} - \beta)'\Big) \\
&= \mathrm{E}\Big((X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'u((X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'u)'\Big) \\
&= \mathrm{E}\Big((X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'uu'Z(Z'\Omega Z)^{-1}Z'X(X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}\Big) \\
&\approx (X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}X'Z(Z'\Omega Z)^{-1}Z'\mathrm{E}(uu')Z(Z'\Omega Z)^{-1}Z'X(X'Z(Z'\Omega Z)^{-1}Z'X)^{-1} \\
&= \sigma^2(X'Z(Z'\Omega Z)^{-1}Z'X)^{-1}.
\end{aligned}
$$

Note that $\beta_{GMM} \longrightarrow \beta$ implies $\mathrm{E}(\beta_{GMM}) \longrightarrow \beta$ in the 1st line.

$\approx$ in the 4th line indicates that Z and X are treated as exogenous variables although they are stochastic.

We assume that $\mathrm{E}(uu') = \sigma^2\Omega$ from the 4th line to the 5th line.

- We derive the asymptotic distribution of β_{GMM}.

From the central limit theorem,

$$\frac{1}{\sqrt{n}} Z'u \longrightarrow N(0,\ \sigma^2 M_{z\Omega z}).$$

Accordingly, β_{GMM} is asymptotically distributed as:

$$\begin{aligned}\sqrt{n}(\beta_{GMM} - \beta) &= \Big((\frac{1}{n}X'Z)(\frac{1}{n}Z'\Omega Z)^{-1}(\frac{1}{n}Z'X) \Big)^{-1} (\frac{1}{n}X'Z)(\frac{1}{n}Z'\Omega Z)^{-1}(\frac{1}{\sqrt{n}}Z'u) \\ &\longrightarrow N(0,\ \sigma^2 (M_{xz} M_{z\Omega z}^{-1} M'_{xz})^{-1}).\end{aligned}$$

Practically, we use: $\beta_{GMM} \sim N\Big(\beta,\ s^2 (X'Z(Z'\Omega Z)^{-1} Z'X)^{-1}\Big)$,

where $s^2 = \dfrac{1}{n-k}(y - X\beta_{GMM})'\Omega^{-1}(y - X\beta_{GMM})$.

We may use n instead of $n - k$.

Identically and Independently Distributed Errors:

- If $u_1, u_2, \cdots, u_n$ are mutually independent and u_i is distributed with mean zero and variance σ^2, the mean and variance of u^* are given by:

$$\mathrm{E}(u^*) = 0 \quad \text{and} \quad \mathrm{V}(u^*) = \mathrm{E}(u^* u^{*\prime}) = \sigma^2 Z'Z.$$

Using GLS, GMM is obtained as:

$$\beta_{GMM} = (X^{*\prime}(Z'Z)^{-1}X^*)^{-1}X^{*\prime}(Z'Z)^{-1}y^* = \Big(X'Z(Z'Z)^{-1}Z'X\Big)^{-1}X'Z(Z'Z)^{-1}Z'y.$$

- We derive the asymptotic distribution of β_{GMM}.

From the central limit theorem,

$$\frac{1}{\sqrt{n}}Z'u \longrightarrow N(0, \sigma^2 M_{zz}).$$

Accordingly, β_{GMM} is distributed as:

$$\sqrt{n}(\beta_{GMM} - \beta) = \Big((\frac{1}{n}X'Z)(\frac{1}{n}Z'Z)^{-1}(\frac{1}{n}Z'X)\Big)^{-1}(\frac{1}{n}X'Z)(\frac{1}{n}Z'Z)^{-1}(\frac{1}{\sqrt{n}}Z'u)$$

$$\longrightarrow\ N\Big(0,\ \sigma^2(M_{xz}M_{zz}^{-1}M_{xz}')^{-1}\Big).$$

Practically, for large n we use the following distribution:

$$\beta_{GMM} \sim N\Big(\beta, s^2(X'Z(Z'Z)^{-1}Z'X)^{-1}\Big),$$

where $s^2 = \dfrac{1}{n-k}(y - X\beta_{GMM})'(y - X\beta_{GMM})$.

- The above GMM is equivalent to 2SLS.

X: $n \times k$, $\quad Z$: $n \times r$, $\quad r > k$.

Assume:

$$\frac{1}{n}X'u = \frac{1}{n}\sum_{i=1}^{n} x_i' u_i \longrightarrow \mathrm{E}(x'u) \neq 0,$$
$$\frac{1}{n}Z'u = \frac{1}{n}\sum_{i=1}^{n} z_i' u_i \longrightarrow \mathrm{E}(z'u) = 0.$$

Regress X on Z, i.e., $X = Z\Gamma + V$ by OLS, where Γ is a $r \times k$ unknown parameter matrix and V is an error term,

Denote the predicted value of X by $\hat{X} = Z\hat{\Gamma} = Z(Z'Z)^{-1}Z'X$, where $\hat{\Gamma} = (Z'Z)^{-1}Z'X$.

Review — IV estimator: Consider the regression model is:

$$y = X\beta + u,$$

Assumption: $\mathrm{E}(X'u) \neq 0$ and $\mathrm{E}(Z'u) = 0$.

The $n \times k$ matrix Z is called the instrumental variable (IV).

The IV estimator is given by:

$$\beta_{IV} = (Z'X)^{-1}Z'y,$$

- Note that 2SLS is equivalent to IV in the case of $Z = \hat{X}$, where this Z is different from the previous Z.

This Z is a $n \times k$ matrix, while the previous Z is a $n \times r$ matrix.

Z in the IV estimator is replaced by $\hat{X}$.

Then,

$$\beta_{2SLS} = (\hat{X}'X)^{-1}\hat{X}'y = \left(X'Z(Z'Z)^{-1}Z'X\right)^{-1}X'Z(Z'Z)^{-1}Z'y = \beta_{GMM}.$$

GMM is interpreted as the GLS applied to MM.