

S is estimated as:

$$\hat{S} = \hat{\Gamma}_0 + \sum_{i=1}^q \left(1 - \frac{i}{q+1}\right) (\hat{\Gamma}_i + \hat{\Gamma}'_i),$$

\implies the Newey-West Estimator

Note that $\hat{S} \rightarrow S$, because $\hat{\Gamma}_\tau \rightarrow \Gamma_\tau$ as $n \rightarrow \infty$.

Asymptotic Properties of GMM:

GMM is consistent and asymptotic normal as follows:

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0, (D'S^{-1}D)^{-1}),$$

where D is a $r \times k$ matrix, and \hat{D} is an estimator of D , defined as:

$$D = \frac{\partial g(\theta; W)}{\partial \theta'}, \quad \hat{D} = \frac{\partial g(\hat{\theta}; W)}{\partial \theta'}.$$

Proof of Asymptotic Normality:

Assumption 1: $\hat{\theta} \rightarrow \theta$

Assumption 2: $\sqrt{ng}(\theta; W) \rightarrow N(0, S)$, i.e., $S = \lim_{n \rightarrow \infty} V(\sqrt{ng}(\theta; W))$.

The first-order condition of GMM is:

$$\frac{\partial g(\theta; W)'}{\partial \theta} S^{-1} g(\theta; W) = 0.$$

The GMM estimator, denote by $\hat{\theta}$, satisfies the above equation.

Therefore, we have the following:

$$\frac{\partial g(\hat{\theta}; W)'}{\partial \theta} \hat{S}^{-1} g(\hat{\theta}; W) = 0.$$

Linearize $g(\hat{\theta}; W)$ around $\hat{\theta} = \theta$ as follows:

$$g(\hat{\theta}; W) = g(\theta; W) + \frac{\partial g(\bar{\theta}; W)}{\partial \theta'}(\hat{\theta} - \theta) = g(\theta; W) + \bar{D}(\hat{\theta} - \theta),$$

where $\bar{D} = \frac{\partial g(\bar{\theta}; W)}{\partial \theta'}$, and $\bar{\theta}$ is between $\hat{\theta}$ and θ .

⇒ **Theorem of Mean Value** (平均値の定理)

Substituting the linear approximation at $\hat{\theta} = \theta$, we obtain:

$$\begin{aligned} 0 &= \hat{D}'\hat{S}^{-1}g(\hat{\theta}; W) \\ &= \hat{D}'\hat{S}^{-1}\left(g(\theta; W) + \bar{D}(\hat{\theta} - \theta)\right) \\ &= \hat{D}'\hat{S}^{-1}g(\theta; W) + \hat{D}'\hat{S}^{-1}\bar{D}(\hat{\theta} - \theta), \end{aligned}$$

which can be rewritten as:

$$\hat{\theta} - \theta = -(\hat{D}'\hat{S}^{-1}\bar{D})^{-1}\hat{D}'\hat{S}^{-1}g(\theta; W).$$

Note that $\bar{D} = \frac{\partial g(\bar{\theta}; W)}{\partial \theta'}$, where $\bar{\theta}$ is between $\hat{\theta}$ and θ .

From Assumption 1, $\hat{\theta} \rightarrow \theta$ implies $\bar{\theta} \rightarrow \theta$

Therefore,

$$\sqrt{n}(\hat{\theta} - \theta) = -(\hat{D}'\hat{S}^{-1}\bar{D})^{-1}\hat{D}'S^{-1} \times \sqrt{ng}(\theta; W).$$

Accordingly , the GMM estimator $\hat{\theta}$ has the following asymptotic distribution:

$$\sqrt{n}(\hat{\theta} - \theta) \rightarrow N(0, (D'S^{-1}D)^{-1}).$$

Note that $\hat{D} \rightarrow D$, $\bar{D} \rightarrow D$, $\hat{S} \rightarrow S$ and Assumption 2 are utilized.

Computational Procedure:

(1) Compute $\hat{S}^{(i)} = \hat{\Gamma}_0 + \sum_{i=1}^q \left(1 - \frac{i}{q+1}\right) (\hat{\Gamma}_i + \hat{\Gamma}'_i)$, where $\hat{\Gamma}_\tau = \frac{1}{n} \sum_{i=\tau+1}^n h(\hat{\theta}; w_i) h(\hat{\theta}; w_{i-\tau})'$.
 q is set by a researcher.

(2) Use the following iterative procedure:

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - (\hat{D}^{(i)'} \hat{S}^{(i-1)} \hat{D}^{(i)})^{-1} \hat{D}^{(i)'} \hat{S}^{(i-1)} g(\hat{\theta}^{(i)}; W).$$

(3) Repeat (1) and (2) until $\hat{\theta}^{(i+1)}$ is equal to $\hat{\theta}^{(i)}$.

In (2), remember that when S is given we take the following iterative procedure:

$$\hat{\theta}^{(i+1)} = \hat{\theta}^{(i)} - (\hat{D}^{(i)'} S^{-1} \hat{D}^{(i)})^{-1} \hat{D}^{(i)'} S^{-1} g(\hat{\theta}^{(i)}; W),$$

where $\hat{D}^{(i)} = \frac{\partial g(\hat{\theta}^{(i)}; W)}{\partial \theta'}$. S is replaced by $\hat{S}^{(i)}$.

- If the assumption $E(h(\theta; w)) = 0$ is violated, the GMM estimator $\hat{\theta}$ is no longer consistent.

Therefore, we need to check if $E(h(\theta; w)) = 0$.

From Assumption 2, note as follows:

$$J = \left(\sqrt{ng}(\hat{\theta}; W) \right)' \hat{S}^{-1} \left(\sqrt{ng}(\hat{\theta}; W) \right) \longrightarrow \chi^2(r - k),$$

which is called Hansen's J test.

Because of r equations and k parameters, the degree of freedom is given by $r - k$.

If J is small enough, we can judge that the specified model is correct.

Testing Hypothesis:

Remember that the GMM estimator $\hat{\theta}$ has the following asymptotic distribution:

$$\sqrt{n}(\hat{\theta} - \theta) \longrightarrow N(0, (D'S^{-1}D)^{-1}).$$

Consider testing the following null and alternative hypotheses:

- The null hypothesis: $H_0 : R(\theta) = 0$,
- The alternative hypothesis: $H_1 : R(\theta) \neq 0$,

where $R(\theta)$ is a $p \times 1$ vector function for $p \leq k$.

p denotes the number of restrictions.

$R(\theta)$ is linearized as: $R(\hat{\theta}) = R(\theta) + R_{\bar{\theta}}(\hat{\theta} - \theta)$, where $R_{\bar{\theta}} = \frac{\partial R(\bar{\theta})}{\partial \theta'}$, which is a $p \times k$ matrix.

Note that $\bar{\theta}$ is between $\hat{\theta}$ and θ . If $\hat{\theta} \rightarrow \theta$, then $\bar{\theta} \rightarrow \theta$ and $R_{\bar{\theta}} \rightarrow R_{\theta}$.

Under the null hypothesis $R(\theta) = 0$, we have $R(\hat{\theta}) = R_{\bar{\theta}}(\hat{\theta} - \theta)$, which implies that the distribution of $R(\hat{\theta})$ is equivalent to that of $R_{\bar{\theta}}(\hat{\theta} - \theta)$.

The distribution of $\sqrt{n}R(\hat{\theta})$ is given by:

$$\sqrt{n}R(\hat{\theta}) = \sqrt{n}R_{\bar{\theta}}(\hat{\theta} - \theta) \rightarrow N(0, R_{\theta}(D'S^{-1}D)^{-1}R'_{\theta}).$$

Therefore, under the null hypothesis, we have the following distribution:

$$nR(\hat{\theta})(R_{\theta}(D'S^{-1}D)^{-1}R'_{\theta})^{-1}R(\hat{\theta})' \rightarrow \chi^2(p).$$

Practically, replacing θ by $\hat{\theta}$ in R_{θ} , D and S , we use the following test statistic:

$$nR(\hat{\theta})(R_{\hat{\theta}}(\hat{D}'\hat{S}^{-1}\hat{D})^{-1}R'_{\hat{\theta}})^{-1}R(\hat{\theta})' \rightarrow \chi^2(p).$$

\Rightarrow Wald type test

Examples of $h(\theta; w)$:

1. OLS:

Regression Model: $y_i = x_i\beta + \epsilon_i$, $E(x_i'\epsilon_i) = 0$

$h(\theta; w_i)$ is taken as:

$$h(\theta; w_i) = x_i'(y_i - x_i\beta).$$

2. IV (Instrumental Variable, 操作变数法):

Regression Model: $y_i = x_i\beta + \epsilon_i$, $E(x_i'\epsilon_i) \neq 0$, $E(z_i'\epsilon_i) = 0$

$h(\theta; w_i)$ is taken as:

$$h(\theta; w_i) = z_i'(y_i - x_i\beta),$$

where z_i is a vector of instrumental variables.

When z_i is a $1 \times k$ vector, the GMM of β is equivalent to the instrumental variable (IV) estimator.

When z_i is a $1 \times r$ vector for $r > k$, the GMM of β is equivalent to the two-stage least squares (2SLS) estimator.

3. **NLS (Nonlinear Least Squares, 非線形最小二乘法):**

Regression Model: $f(y_i, x_i, \beta) = \epsilon_i$, $E(x_i' \epsilon_i) \neq 0$, $E(z_i' \epsilon_i) = 0$

$h(\theta; w_i)$ is taken as:

$$h(\theta; w_i) = z_i' f(y_i, x_i, \beta)$$

where z_i is a vector of instrumental variables.

Example: Demand function using STATA

二人以上の世帯のうち勤労者世帯（全国）

year
y = 実収入（一月当たり，実質データ）
q1 = 穀類支出額（一年当たり，実質データ）
p1 = 穀類価格（相対価格 = 穀類 CPI / 総合 CPI）
p2 = 魚介類価格（相対価格 = 魚介類 CPI / 総合 CPI）
p3 = 肉類価格（相対価格 = 肉類 CPI / 総合 CPI）

year	y	q1	p1	p2	p3
2000	567865	7087.0	1.043390	0.884965	0.818365
2001	561722	6993.1	1.032520	0.886179	0.822154
2002	553768	6934.4	1.031800	0.891282	0.834872
2003	539928	6816.8	1.050410	0.876543	0.843621
2004	547006	6651.6	1.089510	0.865226	0.868313
2005	541367	6615.8	1.020640	0.862745	0.887513
2006	540863	6523.7	1.000000	0.878601	0.891975
2007	543994	6680.5	0.994856	0.886831	0.908436
2008	541821	6494.7	1.043610	0.894523	0.932049
2009	533154	6477.3	1.066870	0.898148	0.934156
2010	539577	6458.2	1.040420	0.889119	0.924352
2011	529750	6448.4	1.025960	0.894081	0.925234
2012	538988	6377.6	1.057170	0.904366	0.917879
2013	542018	6360.7	1.047620	0.909938	0.916149
2014	523953	6174.6	1.016130	0.971774	0.960685
2015	525669	6268.0	1.000000	1.000000	1.000000
2016	527501	6244.8	1.018020	1.019020	1.017020

2017 531693 6106.6 1.027890 1.066730 1.025900

```
. tsset year
   time variable: year, 2000 to 2017
      delta: 1 unit
```

```
. reg q1 y p1 p2 p3 if year>2000.5
```

Source	SS	df	MS	Number of obs	=	17
Model	913640.443	4	228410.111	F(4, 12)	=	25.83
Residual	106100.077	12	8841.67308	Prob > F	=	0.0000
				R-squared	=	0.8960
				Adj R-squared	=	0.8613
Total	1019740.52	16	63733.7825	Root MSE	=	94.03

q1	Coef.	Std. Err.	t	P> t	[95% Conf. Interval]	
y	.0067843	.0045443	1.49	0.161	-.003117	.0166856
p1	-1128.834	998.7698	-1.13	0.280	-3304.966	1047.299
p2	356.8095	806.2301	0.44	0.666	-1399.815	2113.434
p3	-3442.221	1130.078	-3.05	0.010	-5904.448	-979.9931
_cons	6850.563	3179.316	2.15	0.052	-76.57278	13777.7

```
. gmm (q1-{b0}-{b1}*y-{b2}*p1-{b3}*p2-{b4}*p3) if year>2000.5, instruments(y p1
> p2 p3)
```

Step 1

Iteration 0: GMM criterion $Q(b) = 42400764$
 Iteration 1: GMM criterion $Q(b) = 6.781e-12$
 Iteration 2: GMM criterion $Q(b) = 6.781e-12$ (backed up)

Step 2

Iteration 0: GMM criterion $Q(b) = 1.966e-15$
 Iteration 1: GMM criterion $Q(b) = 1.963e-15$ (backed up)
 convergence not achieved

The Gauss-Newton stopping criterion has been met but missing standard errors indicate some of the parameters are not identified.

GMM estimation

Number of parameters = 5
 Number of moments = 5
 Initial weight matrix: Unadjusted Number of obs = 17
 GMM weight matrix: Robust

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
/b0	6850.563	17645.71	0.39	0.698	-27734.4	41435.53
/b1	.0067843	.0282325	0.24	0.810	-.0485504	.062119
/b2	-1128.834	1057.915	-1.07	0.286	-3202.309	944.6415
/b3	356.8095	1565.86	0.23	0.820	-2712.219	3425.838
/b4	-3442.221	5085.561	-0.68	0.498	-13409.74	6525.296

Instruments for equation 1: y p1 p2 p3 _cons

Warning: convergence not achieved

```
. gmm (q1-{b0}-{b1}*y-{b2}*p1-{b3}*p2-{b4}*p3) if year>2000.5, instruments(p1 p2  
> p3 l.p1 l.p2 l.p3)
```

Step 1

Iteration 0: GMM criterion Q(b) = 42404066

Iteration 1: GMM criterion Q(b) = 2790.3146

Iteration 2: GMM criterion Q(b) = 2790.3146

Step 2

Iteration 0: GMM criterion Q(b) = .3201826

Iteration 1: GMM criterion Q(b) = .2469289

Iteration 2: GMM criterion Q(b) = .2469289

GMM estimation

Number of parameters = 5

Number of moments = 7

Initial weight matrix: Unadjusted

Number of obs = 17

GMM weight matrix: Robust

	Coef.	Robust Std. Err.	z	P> z	[95% Conf. Interval]	
/b0	-1192.466	4669.012	-0.26	0.798	-10343.56	7958.63
/b1	.0186312	.0067682	2.75	0.006	.0053657	.0318967

/b2	-1016.864	780.979	-1.30	0.193	-2547.554	513.8271
/b3	-905.5585	598.0885	-1.51	0.130	-2077.79	266.6734
/b4	-499.8064	1147.985	-0.44	0.663	-2749.815	1750.202

Instruments for equation 1: p1 p2 p3 L.p1 L.p2 L.p3 _cons

. estat overid

Test of overidentifying restriction:

Hansen's J chi2(2) = 4.19779 (p = 0.1226)