Video 37: Generalization of GM theorem

Scribe: Arghya Sarkar
bs2021

August 2022

Introduction

In this lecture we will state and proof the general version of the GM theorem. For start we state the form of GM theorem that we used before.

Theorem 1 (Gauss-Markov theorem)**.** *Assume the setup for linear regression. Assume X^TX is non-singular. Then for all $\vec{\mathbf{c}}$, $\vec{\mathbf{c}}'\hat{\vec{\beta}}$ is the 'unique' best linear unbiased estimator (BLUE) of $\vec{\mathbf{c}}'\vec{\beta}$.*

But we want to remove the condition that X^TX being non-singular. This will let us work with all possible X. But in this case we cannot work with all $\vec{c}$ anymore. We need to put some restriction on $\vec{c}$.

Question: What restriction on $\vec{c}$ will allow us to generalize GM theorem? From previous videos we know for $\vec{c}'\hat{\vec{\beta}}$ to be estimable we need $\vec{c} \in \mathcal{R}(X)$. This should be a minimal requirement because without it estimating $\vec{c}'\hat{\vec{\beta}}$ will be meaningless. However, it turns out that this restriction is enough and the following theorem is true.

Theorem 2. *Assume the setup for linear regression. Then for all $\vec{\mathbf{c}} \in \mathcal{R}(X)$, $\vec{\mathbf{c}}'\hat{\vec{\beta}}$ is the 'unique' best linear unbiased estimator (BLUE) of $\vec{\mathbf{c}}'\vec{\beta}$.*

The idea for proof of this theorem is similar to Theorem 1. But there is a problem: since we don't assume X^TX is non-singular so we have multiple solution for $\hat{\vec{\beta}}$. Also the identity $\hat{\vec{\beta}} = (X^TX)^{-1}X^T\vec{y}$ is no longer true. So we need to be careful about that.

Proof. Like before take another unbiased linear estimator $\vec{a}'\vec{y}$ and express it as following

$$\vec{a}'\vec{y} = \vec{\mathbf{c}}'\hat{\vec{\beta}} + (\vec{a}'\vec{y} - \vec{\mathbf{c}}'\hat{\vec{\beta}})$$

So $\mathrm{Var}[\vec{a}'\vec{y}] = \mathrm{Var}(\vec{\mathbf{c}}'\hat{\vec{\beta}}) + \mathrm{Var}(\vec{a}'\vec{y} - \vec{\mathbf{c}}'\hat{\vec{\beta}}) + \mathrm{Cov}(\vec{\mathbf{c}}'\hat{\vec{\beta}}, \vec{a}'\vec{y} - \vec{\mathbf{c}}'\hat{\vec{\beta}})$. Like before we will show that $\mathrm{Cov}(\vec{\mathbf{c}}'\hat{\vec{\beta}}, \vec{a}'\vec{y} - \vec{\mathbf{c}}'\hat{\vec{\beta}}) = 0$. But we cannot find the exact form of

$\mathrm{Cov}(\vec{\mathbf{c'}}\hat{\vec{\beta}}, \vec{a'}\vec{y} - \vec{\mathbf{c'}}\hat{\vec{\beta}})$ like before because we no longer know the exact form of $\hat{\vec{\beta}}$. But now we know $\vec{c} \in \mathcal{R}(X) = \mathcal{C}(X^T X)$. So $\vec{c} = X^T X \vec{\lambda}$ for some $\vec{\lambda}$.
So we see,

$$\vec{c'}\hat{\vec{\beta}} = \vec{\lambda}^T X^T X \hat{\vec{\beta}} = \vec{\lambda}^T X^T \vec{y}$$

Here we used the definition of $\hat{\vec{\beta}}$ that says $\hat{\vec{\beta}}$ is a solution of the equation $X^T X \hat{\vec{\beta}} = X^T \vec{y}$. This is how we can express $\vec{c'}\hat{\vec{\beta}}$ by an expression which does not depend on $\hat{\vec{\beta}}$. Also we know, $\vec{a'}\vec{y}$ is unbiased estimator of $\vec{c'}\vec{\beta}$. So

$$\mathbb{E}[\vec{a'}\vec{y}] = \vec{c'}\vec{\beta} \implies \vec{a'}X\vec{\beta} = \vec{c'}\vec{\beta} \qquad \forall \vec{\beta}$$

Which means

$$\vec{a'}X = \vec{c'}$$

So now we calculate the sought co-varience

$$\begin{aligned}
\mathrm{Cov}(\vec{\mathbf{c'}}\hat{\vec{\beta}}, \vec{a'}\vec{y} - \vec{\mathbf{c'}}\hat{\vec{\beta}}) &= \mathrm{Cov}(\vec{\lambda}^T X^T \vec{y}, (\vec{a'} - \vec{\lambda}^T X^T)\vec{y}) \\
&= (\vec{a'} - \vec{\lambda}^T X^T)\mathrm{Var}(\vec{y}) X \vec{\lambda} \\
&= (\vec{a'} - \vec{\lambda}^T X^T)\sigma^2 I X \vec{\lambda} \\
&= \sigma^2((\vec{a'} - \vec{\lambda}^T X^T) X \vec{\lambda} \\
&= \sigma^2((\vec{a'}X - \vec{\lambda}^T X^T X)\vec{\lambda} \\
&= \sigma^2(\vec{a'}X - \vec{\mathbf{c'}})\vec{\lambda} = 0
\end{aligned}$$

We get $\mathrm{Var}[\vec{a'}\vec{y}] = \mathrm{Var}(\vec{\mathbf{c'}}\hat{\vec{\beta}}) + \mathrm{Var}(\vec{a'}\vec{y} - \vec{\mathbf{c'}}\hat{\vec{\beta}}) \implies \mathrm{Var}[\vec{a'}\vec{y}] \geq \mathrm{Var}(\vec{\mathbf{c'}}\hat{\vec{\beta}})$. Which means $\vec{\mathbf{c'}}\hat{\vec{\beta}}$ has lowest varience among all unbiased linear estimators of $\vec{\mathbf{c'}}\vec{\beta}$, i.e. $\vec{\mathbf{c'}}\hat{\vec{\beta}}$ is a BLUE of $\vec{c'}\vec{\beta}$.
For 'uniqueness' part, note if $\vec{l'}\vec{y}$ is another BLUE of $\vec{\mathbf{c'}}\vec{\beta} \quad \forall \vec{\beta}$. Then we know, $\mathrm{Var}[\vec{l'}\vec{y}] = \mathrm{Var}(\vec{\mathbf{c'}}\hat{\vec{\beta}}) + \mathrm{Var}(\vec{l'}\vec{y} - \vec{\mathbf{c'}}\hat{\vec{\beta}})$. But $\vec{l'}\vec{y}$ and $\vec{\mathbf{c'}}\hat{\vec{\beta}}$ are both BLUE of $\vec{\mathbf{c'}}\vec{\beta}$ so $\mathrm{Var}[\vec{l'}\vec{y}] = \mathrm{Var}(\vec{\mathbf{c'}}\hat{\vec{\beta}})$. So

$$\mathrm{Var}(\vec{l'}\vec{y} - \vec{\mathbf{c'}}\hat{\vec{\beta}}) = 0 \qquad \forall \vec{\beta}$$

Which implies

$$\mathbb{P}_{\vec{\beta}}(\vec{l'}\vec{y} = \vec{\mathbf{c'}}\hat{\vec{\beta}}) = 1 \qquad \forall \vec{\beta}$$

This completes the proof. $\square$