23rd International Workshop on Matrices and Statistics
Faculty of Mathematics and Physics
Ljubljana, Slovenia
June 8–12, 2014

Using the Gram-Schmidt Construction to Develop Linear Models

Lynn R. LaMotte
Louisiana State University Health Sciences Center
New Orleans

GS-plus

$A = (\boldsymbol{a}_1, \ldots, \boldsymbol{a}_m)$

GS, with one little addition, constructs Q, T, such that $Q'Q = \mathrm{I}$, $\mathrm{sp}(Q) = \mathrm{sp}(A)$, and $Q = AT$.

i-th step: $\mathrm{sp}(Q_i) = \mathrm{sp}(\boldsymbol{a}_1, \ldots, \boldsymbol{a}_i)$, $Q_i' Q_i = \mathrm{I}$, $Q_i = AT_i$

$\boldsymbol{a}_{i+1} \in \mathrm{sp}(Q_i)$: $Q_{i+1} = Q_i$, $T_{i+1} = T_i$, $i \longrightarrow i+1$

$\boldsymbol{a}_{i+1} \notin \mathrm{sp}(Q_i)$:

$$
\begin{aligned}
\boldsymbol{q}_{i+1} &= c_{i+1}(\boldsymbol{a}_{i+1} - Q_i Q_i' \boldsymbol{a}_{i+1}) \\
&= c_{i+1} A \boldsymbol{e}_{i+1} - c_{i+1} A T_i Q_i' \boldsymbol{a}_{i+1} \\
&= A \underbrace{c_{i+1}(\boldsymbol{e}_{i+1} - T_i Q_i' \boldsymbol{a}_{i+1})]}_{t_{i+1}} \\
&= A \boldsymbol{t}_{i+1}
\end{aligned}
$$

$Q_{i+1} = (Q_i, \boldsymbol{q}_{i+1})$, $T_{i+1} = (T_i, \boldsymbol{t}_{i+1})$

A sequence of easily-proven propositions leading to:

$\mathcal{S}$ is a linear subspace of $\Re^n \Longrightarrow \exists\ Q$ such that $Q'Q = \mathrm{I}$ and $\mathrm{sp}(Q) = \mathcal{S}$.

Proposition 1

GS establishes that, for any non-zero $n \times m$ matrix A, there exists an $n \times \nu$ matrix Q such that $\mathrm{sp}(Q) = \mathrm{sp}(A)$ and $Q'Q = \mathrm{I}_\nu$.

Proposition 2

If $Q'Q = \mathrm{I}_\nu$ then
(1) $\boldsymbol{z} \in \mathrm{sp}(Q) \iff QQ'\boldsymbol{z} = \boldsymbol{z}$;
and
(2) $\mathrm{sp}(Q) = \mathrm{sp}(QQ')$.

Proposition 3

If Q and R are $n \times \nu$ and $n \times \eta$ matrices, respectively, such that $Q'Q = \mathrm{I}_\nu$ and $R'R = \mathrm{I}_\eta$, then

$$\mathrm{sp}(Q) \subset \mathrm{sp}(R)$$

$$\iff$$

$$RR' - QQ' = (RR' - QQ')(RR' - QQ');$$

and

$$\mathrm{sp}(Q) = \mathrm{sp}(R) \iff RR' = QQ'.$$

Proof uses Prop. 2.

Proposition 4

If $\mathrm{sp}(Q) \subset \mathrm{sp}(R)$ *then* $\eta \geq \nu$. *If* $\mathrm{sp}(Q) = \mathrm{sp}(R)$ *then* $\eta = \nu$.

Proof. $\mathrm{sp}(Q) \subset \mathrm{sp}(R) \Longrightarrow$, by Prop. 3,

$$(RR' - QQ') = (RR' - QQ')(RR' - QQ')$$

$$\Longrightarrow \mathrm{tr}(RR' - QQ') = \eta - \nu \geq 0.$$

$\mathrm{sp}(Q) = \mathrm{sp}(R) \Longrightarrow RR' = QQ'$, by Prop. 3, $\Longrightarrow$
$\mathrm{tr}(RR') = \eta = \mathrm{tr}(QQ') = \nu$. $\square$

Proposition 5

If $\mathrm{sp}(Q) \subset \mathrm{sp}(R)$ *and* $\nu = \eta$ *then* $\mathrm{sp}(Q) = \mathrm{sp}(R)$*.*

Proof. $\mathrm{sp}(Q) \subset \mathrm{sp}(R) \Longrightarrow$, by Prop. 3,

$$\mathrm{tr}[(RR' - QQ')(RR' - QQ')] = \mathrm{tr}(RR' - QQ') = \eta - \nu = 0$$

$\Longrightarrow RR' = QQ'$, $\Longrightarrow$, by Prop. 2(2), $\mathrm{sp}(Q) = \mathrm{sp}(R)$. $\square$

Proposition 6

If $\mathcal{S}$ is a linear subspace of $\Re^n$ and $\mathcal{S} \neq \{\mathbf{0}\}$, then there exists an $n \times \nu$ matrix Q such that $Q'Q = \mathrm{I}_\nu$ and $\mathrm{sp}(Q) = \mathcal{S}$.

Proof is sequential construction of Q_i with ν_i columns, similar to GS. It ends in at most n steps with $\mathrm{sp}(Q_i) = \mathcal{S}$ because $\mathrm{sp}(Q_i) \subset \mathrm{sp}(\mathrm{I}_n)$.

Proposition 7

Let $\mathcal{S}$ be a non-trivial linear subspace of $\Re^n$, and let $\boldsymbol{y}$ be an n-vector. There exists exactly one vector $\hat{\boldsymbol{y}}$ in $\mathcal{S}$ such that $\boldsymbol{y} - \hat{\boldsymbol{y}}$ is in $\mathcal{S}^\perp$.

Proof. $\hat{\boldsymbol{y}} = QQ'\boldsymbol{y} \in \mathrm{sp}(Q) = \mathcal{S}$ and $Q'(\boldsymbol{y} - \hat{\boldsymbol{y}}) = \boldsymbol{0}$.

$\tilde{\boldsymbol{y}} \in \mathcal{S}$ and $Q'(\boldsymbol{y} - \tilde{\boldsymbol{y}}) = \boldsymbol{0} \Longrightarrow QQ'\tilde{\boldsymbol{y}} = \tilde{\boldsymbol{y}}$, by Prop. 2(1),

$\Longrightarrow \boldsymbol{0} = QQ'(\boldsymbol{y} - \tilde{\boldsymbol{y}}) = \hat{\boldsymbol{y}} - \tilde{\boldsymbol{y}}$. □

Proposition 8

Let Q be an $n \times \nu$ matrix such that $Q'Q = \mathrm{I}_\nu$, $\nu \geq 1$.

If P is an $n \times n$ matrix such that, for each $\boldsymbol{y} \in \Re^n$, $P\boldsymbol{y} \in \mathrm{sp}(Q)$ and $\boldsymbol{y} - P\boldsymbol{y} \in \mathrm{sp}(Q)^\perp$, then $P = QQ'$.

Definition 1

The orthogonal projection of $\boldsymbol{y} \in \Re^n$ on a linear subspace $\mathcal{S}$ of $\Re^n$ is the vector $\hat{\boldsymbol{y}} \in \mathcal{S}$ such that $\boldsymbol{y} - \hat{\boldsymbol{y}} \in \mathcal{S}^\perp$.

Let $\mathbf{P}_\mathcal{S}$ denote the matrix such that $\mathbf{P}_\mathcal{S}\boldsymbol{y} \in \mathcal{S}$ and $\boldsymbol{y} - \mathbf{P}_\mathcal{S}\boldsymbol{y} \in \mathcal{S}^\perp \ \forall \ \boldsymbol{y} \in \Re^n$.

Linear equations $A\boldsymbol{x} = \boldsymbol{b}$

GS on $A \longrightarrow Q, T$ such that $\mathrm{sp}(Q) = \mathrm{sp}(A)$, $Q'Q = \mathrm{I}$, $Q = AT$.

Check consistency with $\mathbf{P}_A\boldsymbol{b} - \boldsymbol{b}$.

If $\boldsymbol{b} \in \mathrm{sp}(A)$, then $\boldsymbol{x}_* = TQ'\boldsymbol{b}$ is a solution.

$\{\boldsymbol{x} : A\boldsymbol{x} = \boldsymbol{b}\} = \{\boldsymbol{x}_*\} + \mathrm{sp}(\mathrm{I} - TQ'A)$

GS on $A = (A_1, A_2) \longrightarrow Q = (Q_1, Q_2)$

$\mathrm{sp}(Q_1) = \mathrm{sp}(A_1)$

and $\mathrm{sp}(Q_2) = \mathrm{sp}(A) \cap \mathrm{sp}(A_1)^\perp$.

$Q_2 Q_2' = \mathbf{P}_A - \mathbf{P}_{A_1}$.

Least Squares

Let $\mathcal{S}$ be a linear subspace of $\Re^n$ and let $\boldsymbol{m}_0$ be a given n-vector.

Let $\mathcal{M} = \{\boldsymbol{m}_0\} + \mathcal{S}$.

Proposition 9

Let $\boldsymbol{y} \in \Re^n$.

$$\underset{\boldsymbol{m} \in \mathcal{M}}{\operatorname{argmin}} (\boldsymbol{y} - \boldsymbol{m})'(\boldsymbol{y} - \boldsymbol{m}) = \boldsymbol{m}_0 + \mathbf{P}_{\mathcal{S}}(\boldsymbol{y} - \boldsymbol{m}_0)$$

and

$$\min_{\boldsymbol{m} \in \mathcal{M}} (\boldsymbol{y} - \boldsymbol{m})'(\boldsymbol{y} - \boldsymbol{m}) = (\boldsymbol{y} - \boldsymbol{m}_0)'(\mathrm{I} - \mathbf{P}_{\mathcal{S}})(\boldsymbol{y} - \boldsymbol{m}_0).$$

The Canonical Form of the Regression Model

$$\mathrm{E}(\boldsymbol{Y}) = X\boldsymbol{\beta},\ \mathrm{Var}(\boldsymbol{Y}) = \sigma^2\mathrm{I},$$

with X a given $n \times (k+1)$ matrix and unknown parameters

$$\boldsymbol{\beta} \in \Re^{k+1},\ \sigma^2 > 0.$$

The model for $\boldsymbol{\mu} = \mathrm{E}(\boldsymbol{Y})$ is $\boldsymbol{\mu} \in \mathcal{M} = \mathrm{sp}(X)$.

Given a realization $\boldsymbol{y}$ of $\boldsymbol{Y}$, the objective is inference about the mean vector $\boldsymbol{\mu}$.

Basic LS Statistics

Given $\boldsymbol{y}$, the least-squares (LS) estimate of $\boldsymbol{\mu} = X\beta$ is $\hat{\boldsymbol{y}} = \mathbf{P}_X \boldsymbol{y}$.
A *least-squares solution* for β is $\tilde{\boldsymbol{b}}$ such that $X\tilde{\boldsymbol{b}} = \mathbf{P}_X \boldsymbol{y}$.
From GS on X, $\hat{\boldsymbol{b}} = TQ'\boldsymbol{y}$ is a LS solution.

$SST = \boldsymbol{y}'(\mathrm{I} - \mathbf{P_1})\boldsymbol{y}$
$SSR = \boldsymbol{y}'(\mathbf{P}_X - \mathbf{P_1})\boldsymbol{y}$
$SSE = \boldsymbol{y}'(\mathrm{I} - \mathbf{P}_X)\boldsymbol{y}$
Degrees of freedom
$\nu_R = \mathrm{tr}(\mathbf{P}_X - \mathbf{P_1}), \nu_E = \mathrm{tr}(\mathrm{I} - \mathbf{P}_X)$

Sampling distributions

If $\boldsymbol{Y} \sim \mathbf{N}(\boldsymbol{\mu}, \sigma^2 \mathrm{I})$ and $Q'Q = \mathrm{I}_\nu$

then $\boldsymbol{Z} = (1/\sigma)Q'\boldsymbol{Y} \sim \mathbf{N}((1/\sigma)Q'\boldsymbol{\mu}, \mathrm{I}_\nu)$,

and $(1/\sigma^2)\boldsymbol{Y}'(QQ')\boldsymbol{Y} \sim \chi^2$

with ν degrees of freedom and ncp $\delta^2 = \boldsymbol{\mu}'QQ'\boldsymbol{\mu}/\sigma^2$.

A major teaching objective: develop and justify the test statistic for the general linear hypothesis

$$\mathrm{H}_0 : G'\boldsymbol{\beta} = \boldsymbol{c}_0$$

for given G and $\boldsymbol{c}_0$.

Restricted model

$$\begin{aligned}\mathcal{M}_0 &= \{X\boldsymbol{\beta} \in \mathcal{M} : \boldsymbol{\beta} \in \Re^{k+1} \text{ and } G'\boldsymbol{\beta} = \boldsymbol{c}_0\} \\ &= \{X\boldsymbol{b}_0\} + \mathrm{sp}(XN),\end{aligned}$$

$\boldsymbol{b}_0$: $G'\boldsymbol{b}_0 = \boldsymbol{c}_0$ and $\mathrm{sp}(N) = \mathrm{sp}(G)^{\perp}$.

$$
\begin{aligned}
SSE_{Full} &= \min_{X\beta\in\mathcal{M}} (\boldsymbol{y} - X\beta)'(\boldsymbol{y} - X\beta) \\
&= \boldsymbol{y}'(\mathrm{I} - \mathbf{P}_X)\boldsymbol{y} \\
df_{Full} &= \mathrm{tr}(\mathrm{I} - \mathbf{P}_X) \\
SSE_{Rest} &= \min_{X\beta\in\mathcal{M}_0} (\boldsymbol{y} - X\beta)'(\boldsymbol{y} - X\beta) \\
&= (\boldsymbol{y} - X\boldsymbol{b}_0)'(\mathrm{I} - \mathbf{P}_{XN})(\boldsymbol{y} - X\boldsymbol{b}_0), \\
& \quad G'\boldsymbol{b}_0 = \boldsymbol{c}_0 \text{ and } \mathrm{sp}(N) = \mathrm{sp}(G)^{\perp} \\
df_{Rest} &= \mathrm{sp}(\mathrm{I} - \mathbf{P}_{XN}) \\
\Delta SSE &= SSE_{Rest} - SS_{Full} \\
&= (\boldsymbol{y} - X\boldsymbol{b}_0)'(\mathbf{P}_X - \mathbf{P}_{XN})(\boldsymbol{y} - X\boldsymbol{b}_0), \\
\Delta df &= \mathrm{tr}(\mathbf{P}_X - \mathbf{P}_{XN})
\end{aligned}
$$

The test statistic:

$$\boxed{F = \frac{\Delta SSE/\Delta df}{\hat{\sigma}^2}}$$

with $\hat{\sigma}^2 = SSE_{Full}/df_{Full}$.

Note that $\mathbf{P}_X - \mathbf{P}_{XN}$ can be had as $Q_2 Q_2' = \mathbf{P}_H$ from GS on (XN, X). Then

$$F = \frac{(\boldsymbol{y} - X\boldsymbol{b}_0)'\mathbf{P}_H(\boldsymbol{y} - X\boldsymbol{b}_0)/\mathrm{tr}(\mathbf{P}_H)}{\hat{\sigma}^2}.$$

Bridges

With $\mathrm{sp}(Q) = \mathrm{sp}(A)$ and $Q'Q = \mathrm{I}_\nu$,
column rank$(A) = \dim(\mathrm{sp}(A)) = \nu = \mathrm{tr}(QQ')$.

From GS on A, $A(TQ')A = A$ and $(TQ')A(TQ') = TQ'$:

TQ' is a reflexive generalized inverse of A.
TT' is a reflexive generalized inverse of $A'A$.

Show that the column rank of A equals the row rank of A:

GS on $A \longrightarrow Q, T$; $Q'Q = \mathrm{I}_\nu$, $\mathrm{sp}(Q) = \mathrm{sp}(A)$, $Q = AT$.
The column rank of A is ν.

By Prop. 2(1), $QQ'A = A \Longrightarrow A' = A'QQ' \Longrightarrow$
$\mathrm{sp}(A') \subset \mathrm{sp}(A'Q)$. That also $\mathrm{sp}(A'Q) \subset \mathrm{sp}(A') \Longrightarrow$
$\mathrm{sp}(A') = \mathrm{sp}(A'Q)$.

The ν columns of $A'Q$ are linearly independent:
$A'Q\boldsymbol{z} = \mathbf{0} \Longrightarrow T'A'Q\boldsymbol{z} = Q'Q\boldsymbol{z} = \boldsymbol{z} = \mathbf{0}$.

Therefore the column rank of A' is ν.