Gram-Schmidt and QR Decomposition Example

Suppose that

$$\underset{4\times 3}{\boldsymbol{X}} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 1 & 2 \\ 3 & 2 & 2 \\ 4 & 2 & 1 \end{pmatrix}$$

As on the slides, let

$$\boldsymbol{X}_l = \text{the matrix made of the first } l \text{ columns of } \boldsymbol{X}$$

and consider replacing $\boldsymbol{X}$ with $\underset{4\times 3}{\boldsymbol{Z}}$ having orthogonal columns for which $\mathcal{C}(\boldsymbol{Z}_l) = \mathcal{C}(\boldsymbol{X}_l)$ for all l.

Take as the first column of $\boldsymbol{Z}$

$$\boldsymbol{z}_1 = \boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

Then take

$$\boldsymbol{z}_2 = \boldsymbol{x}_2 - \frac{\langle \boldsymbol{x}_2, \boldsymbol{z}_1 \rangle}{\langle \boldsymbol{z}_1, \boldsymbol{z}_1 \rangle}\boldsymbol{z}_1 = \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix} - \frac{17}{30}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} = \begin{pmatrix} \frac{13}{30} \\ -\frac{4}{30} \\ \frac{9}{30} \\ -\frac{8}{30} \end{pmatrix}$$

Note that

$$\langle \boldsymbol{z}_1, \boldsymbol{z}_2 \rangle = \frac{13 - 8 + 27 - 32}{30} = 0$$

and that $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are both linear combinations of $\boldsymbol{z}_1, \boldsymbol{z}_2$, i.e. $\mathcal{C}(\boldsymbol{Z}_2) = \mathcal{C}(\boldsymbol{X}_2)$.

Then take

$$\boldsymbol{z}_3 = \boldsymbol{x}_3 - \left(\frac{\langle \boldsymbol{x}_3, \boldsymbol{z}_1 \rangle}{\langle \boldsymbol{z}_1, \boldsymbol{z}_1 \rangle}\boldsymbol{z}_1 + \frac{\langle \boldsymbol{x}_3, \boldsymbol{z}_2 \rangle}{\langle \boldsymbol{z}_2, \boldsymbol{z}_2 \rangle}\boldsymbol{z}_2 \right)$$

$$= \begin{pmatrix} 1 \\ 2 \\ 2 \\ 1 \end{pmatrix} - \left(\frac{15}{30}\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix} + \frac{(15/30)}{(330/900)}\begin{pmatrix} \frac{13}{30} \\ -\frac{4}{30} \\ \frac{9}{30} \\ -\frac{8}{30} \end{pmatrix} \right) = \begin{pmatrix} -\frac{2}{22} \\ \frac{26}{22} \\ \frac{2}{22} \\ -\frac{14}{22} \end{pmatrix}$$

It is easy to see that $\langle \boldsymbol{z}_1, \boldsymbol{z}_3 \rangle = 0$ and $\langle \boldsymbol{z}_2, \boldsymbol{z}_3 \rangle = 0$ and with

$$\boldsymbol{Z} = \begin{pmatrix} 1 & \frac{13}{30} & -\frac{2}{22} \\ 2 & -\frac{4}{30} & \frac{26}{22} \\ 3 & \frac{9}{30} & \frac{2}{22} \\ 4 & -\frac{8}{30} & -\frac{14}{22} \end{pmatrix}$$

$\mathcal{C}(\boldsymbol{Z}_3) = \mathcal{C}(\boldsymbol{Z}) = \mathcal{C}(\boldsymbol{X}) = \mathcal{C}(\boldsymbol{X}_3)$.

Notice that for

$$\gamma_{kj} = \begin{cases} 1 & \text{if } j = k \\ \frac{\langle \boldsymbol{z}_k, \boldsymbol{x}_j \rangle}{\langle \boldsymbol{z}_k, \boldsymbol{z}_k \rangle} & \text{if } j > k \\ 0 & \text{otherwise} \end{cases}$$

and

$$\boldsymbol{\Gamma} = (\gamma_{kj}) = \begin{pmatrix} 1 & \frac{17}{30} & \frac{15}{30} \\ 0 & 1 & \frac{15}{11} \\ 0 & 0 & 1 \end{pmatrix}$$

one has

$$\boldsymbol{X} = \begin{pmatrix} 1 & \frac{13}{30} & -\frac{2}{22} \\ 2 & -\frac{4}{30} & \frac{26}{22} \\ 3 & \frac{9}{30} & \frac{2}{22} \\ 4 & -\frac{8}{30} & -\frac{14}{22} \end{pmatrix} \begin{pmatrix} 1 & \frac{17}{30} & \frac{15}{30} \\ 0 & 1 & \frac{15}{11} \\ 0 & 0 & 1 \end{pmatrix} = \boldsymbol{Z\Gamma}$$

Further, with

$$\boldsymbol{D} = \boldsymbol{diag}\left(\langle \boldsymbol{z}_1, \boldsymbol{z}_1 \rangle^{1/2}, \langle \boldsymbol{z}_2, \boldsymbol{z}_2 \rangle^{1/2}, \langle \boldsymbol{z}_3, \boldsymbol{z}_3 \rangle^{1/2} \right) = \boldsymbol{diag}\left(\sqrt{30}, \sqrt{11/30}, \sqrt{20/11} \right)$$

one has

$$\boldsymbol{X} = \boldsymbol{Z}\boldsymbol{D}^{-1}\boldsymbol{D}\boldsymbol{\Gamma} = \boldsymbol{QR}$$

for $\boldsymbol{Q} = \boldsymbol{Z}\boldsymbol{D}^{-1}$ and $\boldsymbol{R} = \boldsymbol{D}\boldsymbol{\Gamma}$. The matix

$$\boldsymbol{Q} = \begin{pmatrix} 1 & \frac{13}{30} & -\frac{2}{22} \\ 2 & -\frac{4}{30} & \frac{26}{22} \\ 3 & \frac{9}{30} & \frac{2}{22} \\ 4 & -\frac{8}{30} & -\frac{14}{22} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{30}} & 0 & 0 \\ 0 & \sqrt{\frac{30}{11}} & 0 \\ 0 & 0 & \sqrt{\frac{11}{20}} \end{pmatrix} = \left(\frac{1}{\sqrt{30}} \boldsymbol{z}_1, \sqrt{\frac{30}{11}} \boldsymbol{z}_2, \sqrt{\frac{11}{20}} \boldsymbol{z}_3 \right)$$

is a version of $\boldsymbol{Z}$ with columns of norm 1 (that thus form an orthonormal basis for $\mathcal{C}(\boldsymbol{X})$).

For any vector of responses/targets $\boldsymbol{Y}$, $\{\boldsymbol{q}_j\}$ the columns of $\boldsymbol{Q}$, and $\widehat{\boldsymbol{Y}}$ the projection of $\boldsymbol{Y}$ onto $\mathcal{C}(\boldsymbol{X})$,

$$\widehat{\boldsymbol{Y}} = \sum_{j=1}^{3} \langle \boldsymbol{Y}, \boldsymbol{q}_j \rangle \boldsymbol{q}_j = \boldsymbol{QQ}'\boldsymbol{Y}$$

since

$$\boldsymbol{Q}'\ \boldsymbol{Y} = \begin{pmatrix} \langle \boldsymbol{Y}, \boldsymbol{q}_1 \rangle \\ \langle \boldsymbol{Y}, \boldsymbol{q}_2 \rangle \\ \langle \boldsymbol{Y}, \boldsymbol{q}_3 \rangle \end{pmatrix}$$

Thus

$$\widehat{\boldsymbol{Y}} = \boldsymbol{QQ}'\boldsymbol{Y} = \boldsymbol{QRR}^{-1}\boldsymbol{Q}'\boldsymbol{Y} = \boldsymbol{X}\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{Y}$$

which means that the ordinary least squares coefficient vector is

$$\widehat{\boldsymbol{\beta}}^{\mathrm{OLS}} = \boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{Y}$$

and the OLS predictor of y for an arbitrary input $\boldsymbol{x}$ is $\widehat{f}(\boldsymbol{x}) = \widehat{\boldsymbol{y}}^{\mathrm{OLS}} = \boldsymbol{x}'\widehat{\boldsymbol{\beta}}^{\mathrm{OLS}} = \boldsymbol{x}'\boldsymbol{R}^{-1}\boldsymbol{Q}'\boldsymbol{Y}$.