Gram-Schmidt Orthonormalization

Gram-Schmidt Orthonormalization:

Suppose $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p$ are LI vectors in $\mathbb{R}^n$.

We seek mutually orthogonal vectors $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ in $\mathbb{R}^n \ni$

$$span\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} = span\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\} \quad \forall\, k = 1, \ldots, p.$$

Define

$$\boldsymbol{U}_0 = \underset{n\times 1}{\boldsymbol{0}} \quad \text{and} \quad \boldsymbol{U}_k = [\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k] \quad k = 1, \ldots, p,$$

where

$$\boldsymbol{u}_k = (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{U}_{k-1}})\boldsymbol{x}_k \quad \forall\, k = 1, \ldots, p.$$

We will show $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$ have the desired properties.

$$\begin{aligned}\boldsymbol{u}_2 &= (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{U}_1})\boldsymbol{x}_2 = (\boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{x}_1})\boldsymbol{x}_2 \\ &= \text{residual vector from the regression of } \boldsymbol{x}_2 \text{ on } \boldsymbol{x}_1.\end{aligned}$$

Likewise, $\boldsymbol{u}_k$ is the residual vector from the regression of $\boldsymbol{x}_k$ on $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{k-1}$ $\forall\, k = 3, \ldots, p$.

(This will follow if we can show $\mathcal{C}(\boldsymbol{U}_k) = span\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \quad \forall\, k = 1, \ldots, p.$)

Now can you show

$$span\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\} = span\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \quad \forall\, k = 1, \ldots, p?$$

Now can you prove mutually orthogonality of $\boldsymbol{u}_1, \ldots, \boldsymbol{u}_p$?

Now let $d_k = \|\boldsymbol{u}_k\| \quad \forall\, k = 1, \ldots, p.$

Define $\boldsymbol{q}_k = \frac{1}{d_k}\boldsymbol{u}_k \quad \forall\, k = 1, \ldots, p.$

Note that

$$\begin{aligned}\boldsymbol{q}_k'\boldsymbol{q}_k &= \frac{1}{d_k^2}\boldsymbol{u}_k'\boldsymbol{u}_k \\ &= \frac{1}{d_k^2}\|\boldsymbol{u}_k\|^2 \\ &= 1 \quad \forall\, k = 1, \ldots, p.\end{aligned}$$

Also $\quad \boldsymbol{q}_i'\boldsymbol{q}_j = \frac{1}{d_i d_j}\boldsymbol{u}_i'\boldsymbol{u}_j = 0 \quad \forall\, i \neq j$.

Thus, $\boldsymbol{q}_1, \ldots, \boldsymbol{q}_p$ are mutually orthonormal.

Furthermore,

$$\begin{aligned} span\{\boldsymbol{q}_1, \ldots, \boldsymbol{q}_k\} &= span\{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_k\} \\ &= span\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\} \quad \forall\, k = 1, \ldots, p. \end{aligned}$$

Show that $\boldsymbol{X} = \boldsymbol{Q}\boldsymbol{R}$, where

$$\boldsymbol{X} \equiv [\boldsymbol{x}_1, \ldots, \boldsymbol{x}_p]$$

$$\boldsymbol{Q} \equiv [\boldsymbol{q}_1, \ldots, \boldsymbol{q}_p] \quad \text{and}$$

$\boldsymbol{R}$ is an upper triangular matrix.

Note that $\boldsymbol{R}$ is unique:

Suppose $\boldsymbol{QR}_1 = \boldsymbol{QR}_2 = \boldsymbol{X}$.

Then $\boldsymbol{Q}'\boldsymbol{QR}_1 = \boldsymbol{Q}'\boldsymbol{QR}_2 \Rightarrow \boldsymbol{R}_1 = \boldsymbol{R}_2$.