Solution (#805) Let X be an $m \times n$ such that $m \geqslant n$ and $X^T X$ is invertible. Let $P = X(X^T X)^{-1} X^T$.
(i) Then by associativity of matrix multiplication

$$\begin{aligned} P^2 &= \left(X(X^T X)^{-1} X^T\right)\left(X(X^T X)^{-1} X^T\right) \\ &= X(X^T X)^{-1}(X^T X)(X^T X)^{-1} X^T \\ &= X(X^T X)^{-1} X^T = P. \end{aligned}$$

(ii) Also by the product rule for transposes and transpose rule for inverses we also have

$$\begin{aligned} P^T &= \left(X(X^T X)^{-1} X^T\right)^T \\ &= X^{TT}\left[(X^T X)^{-1}\right]^T X^T \\ &= X\left[(X^T X)^T\right]^{-1} X^T \\ &= X\left(X^T X^{TT}\right)^{-1} X^T \\ &= X\left(X^T X\right)^{-1} X^T = P. \end{aligned}$$

(iii) By the trace product rule we also have

$$\begin{aligned} \operatorname{trace}(P) &= \operatorname{trace}(X(X^T X)^{-1} X^T) \\ &= \operatorname{trace}(X^T X(X^T X)^{-1}) \\ &= \operatorname{trace}(I_n) = n. \end{aligned}$$

(iv) Further

$$\begin{aligned} X^T(I_m - P) &= X^T\left(I_m - X(X^T X)^{-1} X^T\right) \\ &= X^T - (X^T X)(X^T X)^{-1} X^T \\ &= X^T - X^T = 0. \end{aligned}$$

(v) Say that $\mathbf{y} = X\mathbf{v}$. Then set $\mathbf{v} = (X^T X)^{-1} X^T \mathbf{y}$ and we see

$$X\mathbf{v} = X(X^T X)^{-1} X^T \mathbf{y} = P\mathbf{y} = \mathbf{y}$$

as required.

(vi) The vectors in the column space of X are those of the form $X\mathbf{v}$ for some $\mathbf{v}$ (Proposition 3.125). So (v) shows that $P\mathbf{y} = \mathbf{y}$ for precisely those $\mathbf{y}$ in the column space of X. Further for any $\mathbf{v}$ in $\mathbb{R}_n$ and $\mathbf{y}$ in $\mathbb{R}_m$ we have that

$$\mathbf{v}^T X^T (I_m - P)\mathbf{y} = 0$$

which shows that $\mathbf{y} - P\mathbf{y}$ is perpendicular to every vector $X\mathbf{v}$ in the column space of X.