2. (e) Since $MH = 0$,

$\hat{y} = Hy$ is the projection of y onto $\mathcal{L}$.

(f) By commutativity of the trace operator,

$$tr(H) = trace(X(X^TX)^{-1}X^T)$$
$$= tr(X^TX(X^TX)^{-1})$$
$$= tr(I_d)$$
$$= d = rank(X).$$

3. (a) $XX^T = U\Sigma V^T \cdot V\Sigma^T \cdot U^T$

$= U\Sigma\Sigma^T U^T$ $\quad V^T = V^{-1}$ since V orthogonal

$XX^T \cdot U = U\Sigma\Sigma^T U^T \cdot U$

$= U\Sigma\Sigma^T$ where $\Sigma\Sigma^T$ is diagonal.

for any column vector in U, XX^T rescales the vector, but leaves the direction unchanged.

since $\Sigma\Sigma^T$ is diagnol,

the corresponding eigenvalue is σ_i^2 for the ith column.

similarly,

$X^TXV = V\Sigma^T\Sigma$, the eigenvalue for the ith column is σ_i^2

(b) $XX^T \cdot u_i = \sigma_i^2 u_i$

$X^TX\, v_i = \sigma_i^2 v_i$

$Xv_i = U\Sigma V^T \cdot v_i$

$V^T v_i = e_i$ since V orthogonal.

$\Sigma e_i = \begin{pmatrix} 0 \\ 0 \\ \sigma_i \\ 0 \end{pmatrix}$

thus, $Xv_i = \sigma_i u_i$.

similarly, $X^T u_i = V\Sigma U^T u_i = \sigma_i v_i$.

(c). $\sum_{ij} X_{ij}^2 = tr(XX^T)$

$=$ sum of its eigenvalues

$= \sum_{i=1} \sigma_i^2$

$\|X\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}$

d) $\det(X) = \det(U\Sigma V^T) = \det(U)\cdot\det(\Sigma)\cdot\det(U^T)$

the determinant of an orthogonal matrix is 1

thus $\det(X) = \det(\Sigma)$

$= \prod_{i=1}^{r} \sigma_i = \sigma_1 \cdots \sigma_r$