Rank Deficiency

When deriving $\hat{\beta} = (\mathbf{X}^t\mathbf{X})^{-1}\mathbf{X}^t\mathbf{y}$, we assume the rank of $\mathbf{X}$ is $(p+1)$, so $(\mathbf{X}^t\mathbf{X})^{-1}$ exists.

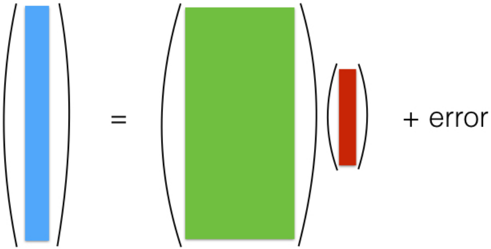
What if $\text{rank}(\mathbf{X}) < p+1$?

$\text{rank}(\mathbf{X}) < p+1$: at least one column of $\mathbf{X}$ is redundant, i.e., it can be reproduced by linear combinations of the other columns.

- X_1: size in sq. ft.; X_2: size in sq. meters;
- X_1: % of population above age 75;

  X_2: % of population below age 18;

  X_3: % of population below between 18 and 75.



Rank Deficiency

- Rank deficiency is not a serious issue: the linear subspace $C(\mathbf{X})$, spanned by the columns of $\mathbf{X}$, is well-defined and therefore $\hat{\mathbf{y}}$ is well-defined and can be computed.
- Due to rank deficiency, $\hat{\beta}$ is not unique.

$$\mathbf{X}_{n\times 2} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ . & . \\ 1 & 2 \end{pmatrix}$$

Rank Deficiency

- Rank deficiency is not a serious issue: the linear subspace $C(\mathbf{X})$, spanned by the columns of $\mathbf{X}$, is well-defined and therefore $\hat{\mathbf{y}}$ is well-defined and can be computed.
- Due to rank deficiency, $\hat{\beta}$ is not unique.
- In `R`, LS coefficients $=$ `NA` means rank deficiency. You can still use the returned model to do prediction.

$$\mathbf{X}_{n\times 2} = \begin{pmatrix} 1 & 2 \\ 1 & 2 \\ . & . \\ 1 & 2 \end{pmatrix}$$