# 1. Least Squares

mtcars =

| | Model | MPG | Cyl | Disp | HP | DRat | WT | QSec | VS |
|---|---|---|---|---|---|---|---|---|---|
| 1 | "Mazda RX4" | 21.0 | 6 | 160.0 | 110 | 3.9 | 2.62 | 16.46 | 0 |
| 2 | "Mazda RX4 Wag" | 21.0 | 6 | 160.0 | 110 | 3.9 | 2.875 | 17.02 | 0 |
| 3 | "Datsun 710" | 22.8 | 4 | 108.0 | 93 | 3.85 | 2.32 | 18.61 | 1 |
| 4 | "Hornet 4 Drive" | 21.4 | 6 | 258.0 | 110 | 3.08 | 3.215 | 19.44 | 1 |
| 5 | "Hornet Sportabout" | 18.7 | 8 | 360.0 | 175 | 3.15 | 3.44 | 17.02 | 0 |
| 6 | "Valiant" | 18.1 | 6 | 225.0 | 105 | 2.76 | 3.46 | 20.22 | 1 |
| 7 | "Duster 360" | 14.3 | 8 | 360.0 | 245 | 3.21 | 3.57 | 15.84 | 0 |
| 8 | "Merc 240D" | 24.4 | 4 | 146.7 | 62 | 3.69 | 3.19 | 20.0 | 1 |
| 9 | "Merc 230" | 22.8 | 4 | 140.8 | 95 | 3.92 | 3.15 | 22.9 | 1 |
| 10 | "Merc 280" | 19.2 | 6 | 167.6 | 123 | 3.92 | 3.44 | 18.3 | 1 |
| ⋮ more | | | | | | | | | |
| 32 | "Volvo 142E" | 21.4 | 4 | 121.0 | 109 | 4.11 | 2.78 | 18.6 | 1 |

$x_1 \cdots x_4$ are the coefficients

intercept term

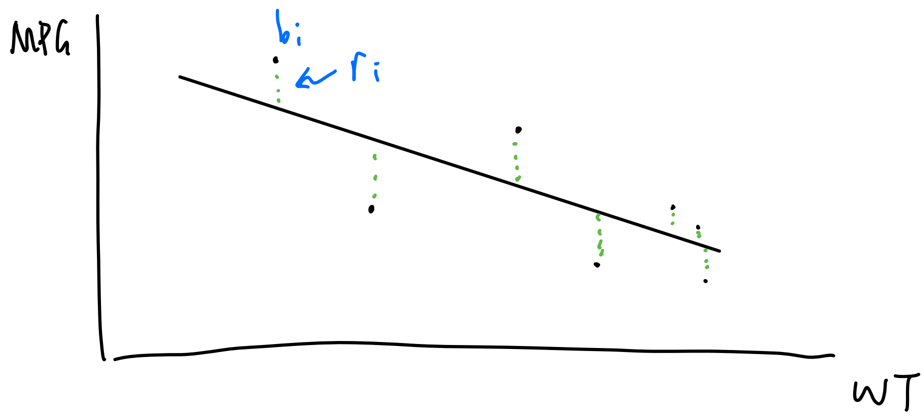
observation

$$MPG_i \sim x_1 + x_2 \cdot Disp_i + x_3 HP_i + x_4 \cdot WT_i \qquad i = 1 \cdots 32$$

$$\min_{x_1 \cdots x_4} \sum_{i}^{32} \left( MPG_i - \left( x_1 + x_2 Disp_i + x_3 HP_i + x_4 WT_i \right) \right)^2$$

## Points of View

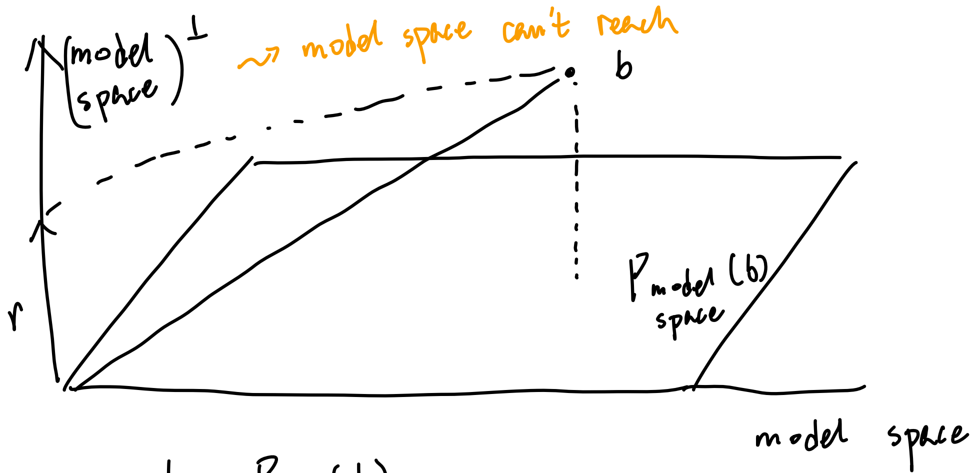
①




↳ choose $x_i$ such that $\sum r_i^2$ $\qquad r_i = b_i - model_i(x)$

## ② Geometry via Linear Algebra

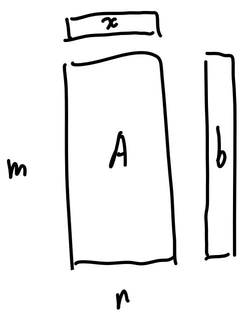
$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix} = \begin{bmatrix} MPG_1 \\ \vdots \\ MPG_m \end{bmatrix} \quad \text{observations}$$



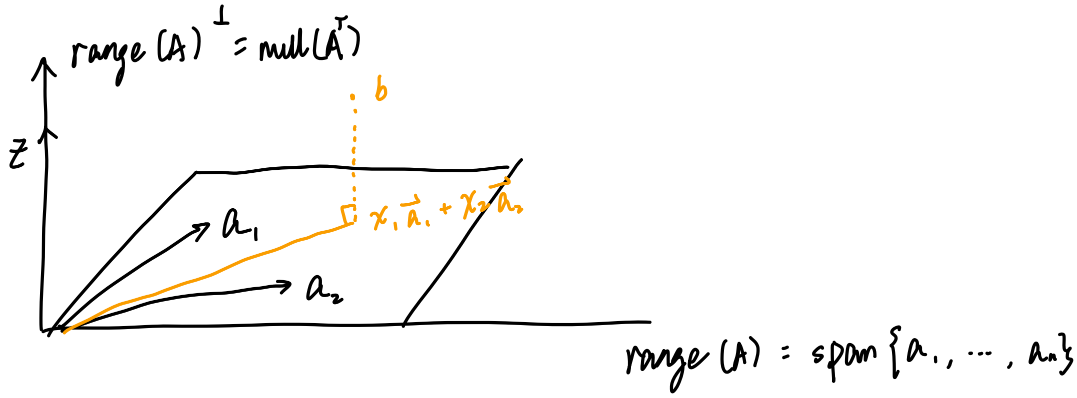

$$r = b - P_{ms}(b)$$

## Matrix Formulation

$$b = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \quad A = \begin{bmatrix} 1 & Disp_1 & HP_1 & WT_1 \\ \vdots & \vdots & \vdots & \vdots \\ 1 & Disp_m & HP_m & WT_m \end{bmatrix} \quad x = \begin{bmatrix} x_1 \\ \vdots \\ x_4 \end{bmatrix}$$




$$\min_x \|r\|^2 \quad \text{such that} \quad \underbrace{Ax}_{\text{linear model}} + \underbrace{r}_{\text{residual}} = b \rightsquigarrow \text{obs}$$




range(A) = { y | Ax = y for some y }

$$0 = z^T y \qquad \forall y \in \text{Range}(A)$$
$$= z^T A x$$
$$= x^T A^T z \qquad \forall x \Rightarrow A^T z = 0$$

$$A = \begin{bmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{bmatrix}$$

$$A^T z = \begin{bmatrix} a_1^T \\ \vdots \\ a_n^T \end{bmatrix} \cdot z = \begin{bmatrix} a_1^T z \\ \vdots \\ a_n^T z \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}$$

$$r \in \text{Null}(A^T) \leftrightarrow A^T r = 0, \quad r = b - A\bar{x}$$
$$A^T(b - A\bar{x}) = 0$$
$$A^T b - A^T A \bar{x} = 0$$
$$A^T A \bar{x} = A^T b$$

Normal Eqn's

to show $\bar{x}$ is unique iff A has full rank

$$p(t) = \alpha + \beta t$$

$$\min_{\alpha,\beta} \sum_{i=1}^{T} (p(t_i) - ppb_i)^2$$

$$A = \begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \qquad b = \begin{bmatrix} ppb_1 \\ \vdots \\ ppb_m \end{bmatrix} \qquad x = \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$$

↳ Todo: $p_n(t) = \alpha + \beta t + \gamma t^2$