Sweep Operator

If we want to solve $Ax=b$ then one way of solving it is work with $[A|b]$ and operate until $[I|x]$

Ex:

$$\left[\begin{array}{cc|c} 2 & 1 & 3 \\ 1 & 3 & 5 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 3/2 \\ 1 & 3 & 5 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 3/2 \\ 0 & 5/2 & 7/2 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1/2 & 3/2 \\ 0 & 1 & 7/5 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{cc|c} 1 & 0 & 4/5 \\ 0 & 1 & 7/5 \end{array}\right]$$ — solution.

We can summarize each row operation with ADJUST.

ADJUST(A, k)

1. $D \leftarrow a_{kk}$; $\text{row}_k \leftarrow \dfrac{\text{row}_k}{D}$

2. for $i \neq k$:

   $B \leftarrow a_{ik}$; $\text{row}_i \leftarrow \text{row}_i - B\,\text{row}_k$

So the solution is given by

$$\text{Adjust}(\cdots(\text{Adjust}([A\ b], 1), 2)\cdots, n)$$

Suppose we want to regress $y \sim X$ which is equivalent to solving

$$X^T X \hat{\beta} = X^T Y$$

Suppose X is $n \times p$, $n > p$, $\text{rank}(X) = p$

So using adjust

$$X^T X \mid X^T y \xrightarrow{\text{Adjust}} I \mid (X^T X)^{-1} X^T Y$$

What about extending

$$\left[\begin{array}{c|c} X^T X & X^T y \\ y^T X & y^T y \end{array}\right] \xrightarrow[\text{rows in } X^T X]{\text{Adjust}} \left[\begin{array}{c|c} I & (X^T X)^{-1} X^T y \\ 0 & y^T y - y^T X (X^T X)^{-1} y \end{array}\right]$$

where

$$
\begin{aligned}
y^T y - y^T X (X^T X)^{-1} y &= y^T (I - H) y \\
&= [(I - H) y]^T [(I - H) y] \\
&= \hat{e}^T \hat{e} = \text{RSS}
\end{aligned}
$$

Extending further

$$\left[\begin{array}{c|cc} X^T X & X^T y & I \\ y^T X & y^T y & 0 \end{array}\right] \xrightarrow[X^T X]{\text{Adust over}} \left[\begin{array}{c|cc} I & \hat{\beta} & (X^T X)^{-1} \\ 0 & \text{RSS} & -\hat{\beta}^T \end{array}\right]$$

From a computational standpoint $\begin{bmatrix} I \\ 0 \end{bmatrix}$ is totally wastefull.

most all inference can be done via this.

Idea: Store $(X^TX)^{-1}, -\hat{\beta}^T$ in place of $\begin{pmatrix} I \\ 0 \end{pmatrix}$. This is what SWEEP does.

SWEEP(A,k)

1. $D \leftarrow a_{kk}$; $\text{row}_k \leftarrow \text{row}_k / D$;

2. for $i \neq k$

   $B \leftarrow a_{ik}$; $\text{row}_i \leftarrow \text{row}_i - B\,\text{row}_k$

   $a_{ik} \leftarrow -B/D$;

3. $a_{kk} \leftarrow 1/D$

Typically used in subset regression.