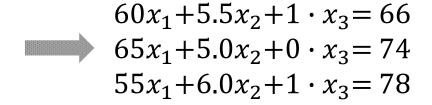
Over-determined Linear Systems (#equations >> #variables)

Recall a linear system:

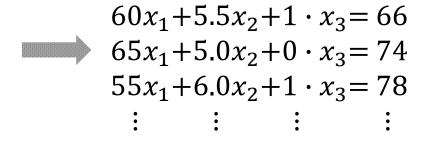
Person ID	Weight	Height	ls_smoking	Life-span
1	60kg	5.5ft	Yes (=1)	66
2	65kg	5.0ft	No (=0)	74
3	55kg	6.0ft	Yes (=1)	78



Over-determined Linear Systems (#equations >> #variables)

- Recall a linear system:
- What if we have much more data examples?

Person ID	Weight	Height	ls_smoking	Life-span
1	60kg	5.5ft	Yes (=1)	66
2	65kg	5.0ft	No (=0)	74
3	55kg	6.0ft	Yes (=1)	78
:	:	:	:	:



• Matrix equation: $\begin{bmatrix} 60 & 5.5 & 1 \\ 65 & 5.0 & 0 \\ 55 & 6.0 & 1 \\ \vdots & \vdots & \vdots \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 66 \\ 74 \\ 78 \\ \vdots \end{bmatrix}$

 $m \gg n$: more equations than variables

Usually no solution exists

Vector Equation Perspective

• Vector equation form: $\begin{bmatrix} 60 \\ 65 \\ 55 \\ \vdots \end{bmatrix} x_1 + \begin{bmatrix} 5.5 \\ 5.0 \\ 6.0 \end{bmatrix} x_2 + \begin{bmatrix} 1 \\ 0 \\ 1 \\ \vdots \end{bmatrix} x_3 = \begin{bmatrix} 66 \\ 74 \\ 78 \\ \vdots \end{bmatrix}$ $\mathbf{a}_1 x_1 + \mathbf{a}_2 x_2 + \mathbf{a}_3 x_3 = \mathbf{b}$

- Compared to the original space \mathbb{R}^n , where $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3, \mathbf{b} \in \mathbb{R}^n$, Span $\{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$ will be a thin hyperplane, so it is likely that $\mathbf{b} \notin \text{Span } \{\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3\}$
 - No solution exists.

Motivation for Least Squares

- Even if no solution exists, we want to approximately obtain the solution for an over-determined system.
- Then, how can we define the best approximate solution for our purpose?



Inner Product

- Given $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, we can consider \mathbf{u} and \mathbf{v} as $n \times 1$ matrices.
- The transpose \mathbf{u}^T is a $1 \times n$ matrix, and the matrix product $\mathbf{u}^T \mathbf{v}$ is a 1×1 matrix, which we write as a scalar without brackets.
- The number $\mathbf{u}^T \mathbf{v}$ is called the **inner product** or **dot product** of \mathbf{u} and \mathbf{v} , and it is written as $\mathbf{u} \cdot \mathbf{v}$.

• For
$$\mathbf{u} = \begin{bmatrix} 3 \\ 2 \\ 1 \end{bmatrix}$$
, $\mathbf{v} = \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix}$, $\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^T \mathbf{v} = \begin{bmatrix} 3 & 2 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 5 \end{bmatrix} = \begin{bmatrix} 14 \end{bmatrix}$

Properties of Inner Product

- Theorem: Let \mathbf{u}, \mathbf{v} , and \mathbf{w} be vectors in \mathbb{R}^n , and let c be a scalar. Then
 - a) $\mathbf{u} \cdot \mathbf{v} = \mathbf{v} \cdot \mathbf{u}$
 - b) $(\mathbf{u} + \mathbf{v}) \cdot \mathbf{w} = \mathbf{u} \cdot \mathbf{w} + \mathbf{v} \cdot \mathbf{w}$
 - c) $(c\mathbf{u}) \cdot \mathbf{v} = c(\mathbf{u} \cdot \mathbf{v}) = \mathbf{u} \cdot (c\mathbf{v})$
 - d) $\mathbf{u} \cdot \mathbf{u} \ge 0$, and $\mathbf{u} \cdot \mathbf{u} = 0$ if and only if $\mathbf{u} = \mathbf{0}$
- Properties (b) and (c) can be combined to produce the following useful rule:

$$(c_1\mathbf{u}_1 + \dots + c_p\mathbf{u}_p) \cdot \mathbf{w} = c_1(\mathbf{u}_1 \cdot \mathbf{w}) + \dots + c_p(\mathbf{u}_p \cdot \mathbf{w})$$