Chapter 2.1 Matrix Operations...

Matrix - Matrix multiplication is composition of functions... works because composition of linear functions is again linear...

linear function

① $f(u+v) = f(u) + f(v)$ $\quad u, v \in \mathbb{R}^n$

② $f(\alpha u) = \alpha f(u)$ $\quad u \in \mathbb{R}^n \quad \alpha \in \mathbb{R}$

Composition Let f and g be linear functions

verification of first property

$$(f \circ g)(u+v) = f(\underbrace{g(u+v)})$$

since g is linear $g(u+v) = g(u) + g(v)$

$$= f(g(u) + g(v)) = f(g(u)) + f(g(v))$$

since f is linear

$$= (f \circ g)(u) + (f \circ g)v ..$$

verification of second property

easier...

vectors of length n

Let $f: \mathbb{R}^n \to \mathbb{R}^m$ ← vectors of length m.

$g: \mathbb{R}^p \to \mathbb{R}^n$

output, input

$f(x) = Ax$ where $A \in \mathbb{R}^{m \times n}$

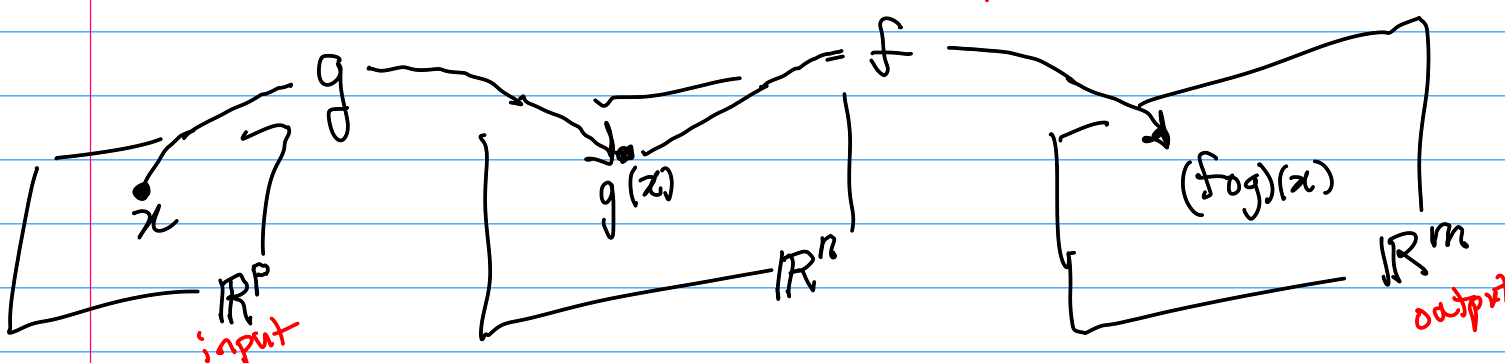
$g(x) = Bx$ where $B \in \mathbb{R}^{n \times p}$ ← columns. & rows

$$(f \circ g)(x) = f(g(x)) = f(Bx) = A(Bx) = (AB)x$$

$m \times n \quad n \times p$ — matches

$AB \in \mathbb{R}^{m \times p}$ (output: m, input: p)

Matrix corresponding to the composition



Example: Goal to find the matrix for fog.

$$f(x) = \begin{bmatrix} 2x_1 - x_2 + 3x_3 \\ -x_1 + 5x_2 - 3x_3 \end{bmatrix} \qquad A = \begin{bmatrix} 2 & -1 & 3 \\ -1 & 5 & -3 \end{bmatrix}$$

$$g(x) = \begin{bmatrix} -x_1 + x_2 \\ x_1 + 2x_2 \\ 2x_1 + x_2 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$(f \circ g)(x) = f(g(x)) = f\left(\begin{bmatrix} -x_1 + x_2 \\ x_1 + 2x_2 \\ 2x_1 + x_2 \end{bmatrix} \right)$$

$$= \begin{bmatrix} 2(-x_1 + x_2) - (x_1 + 2x_2) + 3(2x_1 + x_2) \\ -(-x_1 + x_2) + 5(x_1 + 2x_2) - 3(2x_1 + x_2) \end{bmatrix}$$

Next step... factor out the x's so we can see what the matrix is...

Easier to figure it out using letters

$$A = \begin{bmatrix} 2 & -1 & 3 \\ -1 & 5 & -3 \end{bmatrix} \qquad B = \begin{bmatrix} -1 & 1 \\ 1 & 2 \\ 2 & 1 \end{bmatrix}$$

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \qquad B = \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{bmatrix}$$

row column

$$f(x) = Ax = \begin{bmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{bmatrix}$$

$$g(x) = Bx = \begin{bmatrix} b_{11}x_1 + b_{12}x_2 \\ b_{21}x_1 + b_{22}x_2 \\ b_{31}x_1 + b_{32}x_2 \end{bmatrix}$$

$$(f \circ g)(x) = f\left(\begin{bmatrix} b_{11}x_1 + b_{12}x_2 \\ b_{21}x_1 + b_{22}x_2 \\ b_{31}x_1 + b_{32}x_2 \end{bmatrix} \right)$$

$$= \begin{bmatrix} a_{11}(b_{11}x_1 + b_{12}x_2) + a_{12}(b_{21}x_1 + b_{22}x_2) + a_{13}(b_{31}x_1 + b_{32}x_2) \\ a_{21}(b_{11}x_1 + b_{12}x_2) + a_{22}(b_{21}x_1 + b_{22}x_2) + a_{23}(b_{31}x_1 + b_{32}x_2) \end{bmatrix}$$

$$= \begin{bmatrix} (a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31})x_1 + (a_{11}b_{12} + a_{12}b_{21} + a_{13}b_{31})x_2 \\ (a_{22}b_{11} + a_{22}b_{21} + a_{23}b_{31})x_1 + (a_{22}b_{12} + a_{22}b_{22} + a_{23}b_{32})x_2 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{21} + a_{13}b_{31} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{bmatrix}}_{\text{The matrix for } AB} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

That wasn't easy enough... try again...

$f(x) = Ax \qquad g(x) = Bx$

$(f \circ g)(x) = f(\underbrace{Bx}_{\text{what is } Bx})$

matrix-vector

column representation

column vectors

$$Bx = \begin{bmatrix} b_1 & b_2 \end{bmatrix} x = x_1 b_1 + x_2 b_2$$

scalars

$$(f \circ g)(x) = f(x_1 b_1 + x_2 b_2) = x_1 f(b_1) + x_2 f(b_2)$$

by linearity of f.

matrix-vector mult
use row representation

$$f(b_1) = Ab_1 = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_1 = \begin{bmatrix} a_1 \cdot b_1 \\ a_2 \cdot b_1 \end{bmatrix}$$

$$f(b_2) = Ab_2 = \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} b_2 = \begin{bmatrix} a_1 \cdot b_2 \\ a_2 \cdot b_2 \end{bmatrix}$$

$$(f \circ g)(x) = x_1 \begin{bmatrix} a_1 \cdot b_1 \\ a_2 \cdot b_1 \end{bmatrix} + x_2 \begin{bmatrix} a_1 \cdot b_2 \\ a_2 \cdot b_2 \end{bmatrix}$$

$$= \begin{bmatrix} (a_1 \cdot b_1)x_1 \\ (a_2 \cdot b_1)x_1 \end{bmatrix} + \begin{bmatrix} (a_1 \cdot b_2)x_2 \\ (a_2 \cdot b_2)x_2 \end{bmatrix}$$

$$= \begin{bmatrix} (a_1 \cdot b_1)x_1 + (a_1 \cdot b_2)x_2 \\ (a_2 \cdot b_1)x_1 + (a_2 \cdot b_2)x_2 \end{bmatrix}$$

$$= \underbrace{\begin{bmatrix} a_1 \cdot b_1 & a_1 \cdot b_2 \\ a_2 \cdot b_1 & a_2 \cdot b_2 \end{bmatrix}}_{\text{Matrix for } AB} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

here a_1, a_2 are rows of A

b_1, b_2 were columns of B

In general

$$AB = \begin{bmatrix} a_1 \cdot b_1 & a_1 \cdot b_2 & \cdots & a_1 \cdot b_p \\ \vdots & & & \\ a_m \cdot b_1 & a_m \cdot b_2 & \cdots & a_m \cdot b_p \end{bmatrix} \in \mathbb{R}^{m \times p}$$