# Matrix Vector Product is Linear

Let $\mathbf{x}$ and $\mathbf{y}$ be vectors of length, $N$, and let $\mathbf{A}$ be an $N \times N$ matrix and where:

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

Matrix-vector product is *linear* because it satisfies the following two constraints:

1. $\mathbf{A}(\mathbf{x} + \mathbf{y}) = \mathbf{A}\mathbf{x} + \mathbf{A}\mathbf{y}$
2. $\mathbf{A}(c\mathbf{x}) = c\mathbf{A}\mathbf{x}$

for scalar constant, $c$.

## Matrix Vector Product is Linear (contd.)

It is is easy to show that matrix-vector product satisfies the first constraint:

$$\sum_{j=0}^{N-1} A_{ij}(x_j + y_j) = \sum_{j=0}^{N-1} A_{ij}x_j + \sum_{j=0}^{N-1} A_{ij}y_j$$

...and the second:

$$\sum_{j=0}^{N-1} A_{ij}\, c\, x_j = c \sum_{j=0}^{N-1} A_{ij}x_j.$$

## Kronecker Delta Function

The *Kronecker delta* function is defined as follows:

$$\delta(i) = \begin{cases} 1 & \text{if } i = 0 \\ 0 & \text{otherwise.} \end{cases}$$

## Shift Matrix

$$\mathbf{S}^0 = \delta(i - j - 0 \bmod 4) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

$$\mathbf{S}^1 = \delta(i - j - 1 \bmod 4) = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix}$$

$$\mathbf{S}^2 = \delta(i - j - 2 \bmod 4) = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}$$

$$\mathbf{S}^3 = \delta(i - j - 3 \bmod 4) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix}$$

Shift Matrix (contd.)

$$\mathbf{S}^0\mathbf{x} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix}$$

$$\mathbf{S}^1\mathbf{x} = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 4 \\ 1 \\ 2 \\ 3 \end{bmatrix}$$

$$\mathbf{S}^2\mathbf{x} = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 3 \\ 4 \\ 1 \\ 2 \end{bmatrix}$$

$$\mathbf{S}^3\mathbf{x} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \\ 4 \\ 1 \end{bmatrix}$$

Shift-invariance

If for matrices, $\mathbf{A}$ and $\mathbf{B}$, it is the case that:

$$\mathbf{AB} = \mathbf{BA}$$

then we say that the matrix product *commutes*. This property can be depicted using a *commutative diagram*:

$$\begin{array}{ccc} \mathbf{x} & \xrightarrow{\mathbf{B}} & \mathbf{Bx} \\ \downarrow \mathbf{A} & & \downarrow \mathbf{A} \\ \mathbf{Ax} & \xrightarrow{\mathbf{B}} & \mathbf{y} \end{array}$$

## Shift-invariance (contd.)

What structure must a $3 \times 3$ matrix, $\mathbf{A}$, possess if its product with each of $\mathbf{S}^n$ for $1 \leq n \leq 3$ is to commute?

$$\mathbf{A} = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$$

We observe that because:

$$\mathbf{S}^n = \underbrace{\mathbf{S}^1 \ldots \mathbf{S}^1}_{n}$$

it suffices to show that $\mathbf{A}$ commutes with $\mathbf{S}^1$.

## Shift-invariance (contd.)

Let's look at the effect of multiplying **A** by $\mathbf{S}^{\mathbf{1}}$:

$$\mathbf{AS}^1 = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} b & c & a \\ e & f & d \\ h & i & g \end{bmatrix}$$

...and multiplying $\mathbf{S}^{\mathbf{1}}$ by **A**:

$$\mathbf{S}^1\mathbf{A} = \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = \begin{bmatrix} g & h & i \\ a & b & c \\ d & e & f \end{bmatrix}$$

## Shift-invariance (contd.)

Clearly, $\mathbf{AS}^1 = \mathbf{S}^1\mathbf{A}$, iff:

$$\begin{bmatrix} b & c & a \\ e & f & d \\ h & i & g \end{bmatrix} = \begin{bmatrix} g & h & i \\ a & b & c \\ d & e & f \end{bmatrix}$$

Let's see if we can see any pattern in these equalities:

$$\begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix} \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} \begin{matrix} a & b & c \\ d & e & f \\ g & h & i \end{matrix}$$

## Shift-invariance (contd.)

What structure must an $N \times N$ matrix, $\mathbf{A}$, possess if its product with each of $\mathbf{S}^n$ for $1 \leq n \leq N$ is to commute?

$$A(i, j) = A(i + n \bmod N, j + n \bmod N)$$

$$\mathbf{A} = \begin{bmatrix} a_0 & a_1 & \dots & a_{N-1} \\ a_{N-1} & a_0 & \dots & a_{N-2} \\ \vdots & \vdots & \ddots & \vdots \\ a_1 & a_2 & \dots & a_0 \end{bmatrix}$$

A matrix with the above structure is termed *circulant*.

Shift-invariance (contd.)

$$\mathbf{AS}^0 = \mathbf{S}^0\mathbf{A} = \begin{bmatrix} a & b & c & d \\ d & a & b & c \\ c & d & a & b \\ b & c & d & a \end{bmatrix}$$

$$\mathbf{AS}^1 = \mathbf{S}^1\mathbf{A} = \begin{bmatrix} b & c & d & a \\ a & b & c & d \\ d & a & b & c \\ c & d & a & b \end{bmatrix}$$

$$\mathbf{AS}^2 = \mathbf{S}^2\mathbf{A} = \begin{bmatrix} c & d & a & b \\ b & c & d & a \\ a & b & c & d \\ d & a & b & c \end{bmatrix}$$

$$\mathbf{AS}^3 = \mathbf{S}^3\mathbf{A} = \begin{bmatrix} d & a & b & c \\ c & d & a & b \\ b & c & d & a \\ a & b & c & d \end{bmatrix}$$

## Discrete Convolution

Recall that for a circulant matrix, **A**, the following is true for all $n$:

$$A(i,j) = A(i+n \bmod N, j+n \bmod N).$$

In particular, it is true when $n = -j$:

$$A(i,j) = A(i-j \bmod N, 0).$$

Recall that $\mathbf{y} = \mathbf{Ax}$ can be written:

$$\begin{aligned} y_i &= \sum_{j=0}^{N-1} A(i,j)x_j \\ &= \sum_{j=0}^{N-1} A(i-j \bmod N, 0)x_j. \end{aligned}$$

<u>Discrete Convolution (contd.)</u>

Let

$$h(i-j \bmod N) = A(i-j \bmod N, 0)$$

then

$$y_i = \sum_{j=0}^{N-1} h(i-j \bmod N) x_j.$$

We say that $\mathbf{y}$ is the *discrete periodic convolution* of $\mathbf{h}$ and $\mathbf{x}$:

$$\mathbf{y} = \mathbf{h} * \mathbf{x}.$$

Example

The weight matrix for a 1*D* artificial retina applied to a simple step pattern:

$$\begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{bmatrix}$$

Example (contd.)

The *convolution kernel* is the first column of the weight matrix:

$$\mathbf{h} = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ -1 \end{bmatrix}$$

$$y_i = \sum_{j=0}^{N-1} h(i - j \bmod N) x_j$$

## Operators

An *operator* is a function which takes one or more functions as arguments and returns a function as its value.

- The gradient operator:

$$f \xrightarrow{\nabla} f'$$

where $f'(x) = df(x)/dx$.

- The addition operator:

$$f, g \xrightarrow{+} \{f + g\}$$

where $\{f+g\}(x) = f(x) + g(x)$.

- The convolution operator:

$$f, g \xrightarrow{*} \{f * g\}$$

where $\{f*g\}(x) = \int_{-\infty}^{\infty} f(y)g(x-y)dy$.

## Digression on High Level Languages

One advantage of high level languages (like Scheme) is that you can actually write operators:

```
(define op+
  (lambda (f g)
    (lambda (x)
      (+ (f x) (g x)))))
```

## Linear Operators

Let $f$ and $g$ be functions and let $\mathcal{A}$ be an operator:

$$f \xrightarrow{\mathcal{A}} \mathcal{A}\, f$$
$$g \xrightarrow{\mathcal{A}} \mathcal{A}\, g.$$

An operator is *linear* if and only if:

1. $\mathcal{A}\,\{f+g\} = \mathcal{A}\, f + \mathcal{A}\, g$
2. $\mathcal{A}\,\{c \cdot f\} = c \cdot \mathcal{A}\, f$

for scalar constant, $c$.

## Examples

The gradient operator is linear:

$$
\begin{aligned}
f + g &\xrightarrow{\nabla} f' + g' \\
c \cdot f &\xrightarrow{\nabla} c \cdot f'
\end{aligned}
$$

## Examples (contd.)

The convolution with $h$ operator is linear:

$$f+g \xrightarrow{*h} f*h+g*h$$
$$c\cdot f \xrightarrow{*h} \{c\cdot\{f*h\}\}$$

$$\{\{f+g\}*h\}(x)=\int_{-\infty}^{\infty}[f(y)+g(y)]h(x-y)dy$$
$$=\int_{-\infty}^{\infty}f(y)h(x-y)dy+\int_{-\infty}^{\infty}g(y)h(x-y)dy$$
$$=\{f*h+g*h\}(x)$$

and

$$\{\{c\cdot f\}*h\}(x)=\int_{-\infty}^{\infty}c\ f(y)h(x-y)dy$$
$$=c\int_{-\infty}^{\infty}f(y)h(x-y)dy$$
$$=\{c\cdot\{f*h\}\}(x)$$

# Linear Operators

Linear operator, $\mathcal{A}$ , takes function, $f$, as input and returns function, $g$, as output:

$$g(t) = \int_{-\infty}^{\infty} A(t,\tau) f(\tau) d\tau.$$

It is easy to verify that $\mathcal{A}$ is linear:

$$\int_{-\infty}^{\infty} A(t,\tau)[f(\tau)+g(\tau)]d\tau =$$

$$\int_{-\infty}^{\infty} A(t,\tau) f(\tau) d\tau + \int_{-\infty}^{\infty} A(t,\tau) g(\tau) d\tau$$

and

$$\int_{-\infty}^{\infty} A(t,\tau)\ c\ f(\tau) d\tau =$$

$$c \int_{-\infty}^{\infty} A(t,\tau) f(\tau) d\tau$$

## Shift Operator

The *shift operator*, $\mathcal{S}_\Delta$, takes a function, $f$, as argument, and returns the same function shifted to the right by $\Delta$:

$$f \xrightarrow{\mathcal{S}_\Delta} f_\Delta$$

where $f_\Delta(t) = f(t - \Delta)$. An operator, $\mathcal{A}$, is *shift-invariant*, if and only if it commutes with the shift operator. This property can be depicted using a commutative diagram:

$$\begin{array}{ccc} f & \xrightarrow{\mathcal{S}_\Delta} & f_\Delta \\ \downarrow \mathcal{A} & & \downarrow \mathcal{A} \\ \mathcal{A} f & \xrightarrow{\mathcal{S}_\Delta} & g \end{array}$$

## Linear Shift-invariant Operators

Linear operator, $\mathcal{A}$, takes function, $f$, as input and returns function, $g$, as output:

$$g(t) = \int_{-\infty}^{\infty} A(t,\tau) f(\tau) d\tau.$$

Let $s_\Delta f = f_\Delta$ and $s_\Delta g = g_\Delta$. If $\mathcal{A}$ is shift-invariant, then:

$$g_\Delta(t) = \int_{-\infty}^{\infty} A(t,\tau) f_\Delta(\tau) d\tau.$$

However, $f_\Delta(\tau)$ is just $f(\tau - \Delta)$ and $g_\Delta(t)$ is just $g(t-\Delta)$:

$$g(t-\Delta) = \int_{-\infty}^{\infty} A(t,\tau) f(\tau-\Delta) d\tau.$$

Adding $\Delta$ to $t$ and $\tau$ throughout:

$$g(t) = \int_{-\infty}^{\infty} A(t+\Delta,\tau+\Delta) f(\tau) d\tau.$$

## Linear Shift-invariant Operators (contd.)

We conclude that the following must be true if $\mathcal{A}$ is shift-invariant:

$$A(t,\tau) = A(t+\Delta,\tau+\Delta).$$

Observe that $A(t,\tau)$ is unchanged by adding $\Delta$ to both arguments. This means that the $2D$ function, $A(t,\tau)$, is equal to a $1D$ function of the difference of $t$ and $\tau$:

$$A(t,\tau) = h(t-\tau).$$

Consequently,

$$g(t) = \int_{-\infty}^{\infty} h(t-\tau) f(\tau) d\tau.$$

This is the *convolution integral*. Consequently:

$$g = h * f.$$

## Linear Shift-invariant Operators (contd.)

Deep Thought: All linear shift invariant operators can be represented as convolutions.