NYU COURANT INSTITUTE OF MATHEMATICAL SCIENCES

NYU DATA SCIENCE

Stationarity (blended lecture)

DS-GA 1013 / MATH-GA 2824 Mathematical Tools for Data Science

Carlos Fernandez-Granda

Review questions

Eigendecomposition of continuous linear-translation invariant functions

Complexity of convolutions

Convolutions in probability

Review questions

1. Why do we almost never apply generic linear models to audio, images and other high-dimensional signals?
2. What is the limitation of PCA in uncovering *interesting* structure in audio, images and similar signals?

Review questions

Eigendecomposition of continuous linear-translation invariant functions

Complexity of convolutions

Convolutions in probability

Linear translation-invariant (LTI) function

A function $\mathcal{F}$ from $\mathbb{C}^N$ to $\mathbb{C}^N$ is linear if for any $x, y \in \mathbb{C}^N$ and any $\alpha \in \mathbb{C}$

$$\mathcal{F}(x + y) = \mathcal{F}(x) + \mathcal{F}(y)$$
$$\mathcal{F}(\alpha x) = \alpha \mathcal{F}(x)$$

and translation invariant if for any shift $0 \leq s \leq N - 1$

$$\mathcal{F}(x^{\downarrow s}) = \mathcal{F}(x)^{\downarrow s}$$

What are the eigenvectors of discrete LTI functions?

What are the eigenvectors of discrete LTI functions?

1. Discrete linear-translation invariant function is equivalent to convolution with impulse response

$$\mathcal{F}(x) = x * h$$

2. Convolution is equivalent to taking DFT, weighting DFT coefficients, and applying inverse DFT

$$\begin{aligned}\mathcal{F}(x) &= \frac{1}{N} F_{[N]}^{*} \operatorname{diag}(\hat{h}) F_{[N]} x \\ &= \frac{1}{N} \sum_{k=0}^{N-1} \hat{h}[k] \hat{x}[k] \psi_k\end{aligned}$$

3. This implies

$$\mathcal{F}(\psi_k) = \hat{h}[k] \psi_k$$

Continuous LTI functions

A function or system $\mathcal{F}$ that maps functions in $[0, 1]$ to functions in $[0, 1]$ is linear if for any functions f, g and any $\alpha \in \mathbb{C}$

$$\mathcal{F}(f + g) = \mathcal{F}(f) + \mathcal{F}(g)$$
$$\mathcal{F}(\alpha f) = \alpha \mathcal{F}(f)$$

and translation invariant if for any shift $s \in \mathbb{R}$

$$\mathcal{F}(f^{\downarrow s}) = \mathcal{F}(f)^{\downarrow s}$$

where $f^{\downarrow s}(t) = f(t - s)$ for all $t \in [0, 1]$

Are complex exponentials eigenvectors of continuous LTI functions?

$$\phi_k(t) := \exp(i2\pi kt)$$

$$y := \mathcal{F}(\phi_k) \ ?$$

Are complex exponentials eigenvectors of continuous LTI functions?

$$\phi_k(t) := \exp(i2\pi kt)$$

$$y := \mathcal{F}(\phi_k) \ ?$$

$$\begin{aligned}
\mathcal{F}(\phi_k^{\downarrow s}) &= y^{\downarrow s} \\
\mathcal{F}(\phi_k^{\downarrow s}) &= \mathcal{F}(\exp(-i2\pi ks)\phi_k^{\downarrow s}) \\
&= \exp(-i2\pi ks)y
\end{aligned}$$

$$y(a-b) = \exp(-i2\pi kb)y(a)$$

$$y(t) = y(0)\exp(-i2\pi kt) \quad \text{Yes!}$$

Review questions

Eigendecomposition of continuous linear-translation invariant functions

Complexity of convolutions

Convolutions in probability

Complexity of convolutions

- Complexity of circular convolution between N-dimensional vectors?

$$x * y\,[j] := \sum_{s=0}^{N-1} x\,[s]\, y^{\downarrow s}\,[j]\,, \quad 0 \leq j \leq N-1$$

- Can we improve it?

Complexity of convolutions

- Complexity of circular convolution between N-dimensional vectors? $\mathcal{O}(N^2)$

$$x * y\,[j] := \sum_{s=0}^{N-1} x\,[s]\, y^{\downarrow s}\,[j]\,, \quad 0 \leq j \leq N-1$$

- Can we improve it? Yes, to $\mathcal{O}(N \log N)$ by using FFTs

Complexity of convolutions

- Complexity of convolution between vectors of dimension M and N?

$$x * y\,[j] := \sum_{s=0}^{M-1} x\,[s]\, y^{\downarrow s}\,[j]\,, \quad 0 \leq j \leq N-1$$

- In convolutional neural networks, we don't use FFTs. Why?

Complexity of convolutions

- Complexity of convolution between vectors of dimension M and N?
  $\mathcal{O}(MN)$

$$x * y\,[j] := \sum_{s=0}^{M-1} x\,[s]\, y^{\downarrow s}\,[j]\,, \quad 0 \leq j \leq N-1$$

- In convolutional neural networks, we don't use FFTs. Why?
  M is very small

Review questions

Eigendecomposition of continuous linear-translation invariant functions

Complexity of convolutions

Convolutions in probability

Sum of independent random variables

Let $\tilde{x}$ and $\tilde{y}$ be independent discrete random variables with N possible values

What is the distribution of $\tilde{z} := \tilde{x} + \tilde{y}$

$$p_{\tilde{z}}(z) =$$

Sum of independent random variables

Let $\tilde{x}$ and $\tilde{y}$ be independent discrete random variables with N possible values

What is the distribution of $\tilde{z} := \tilde{x} + \tilde{y}$

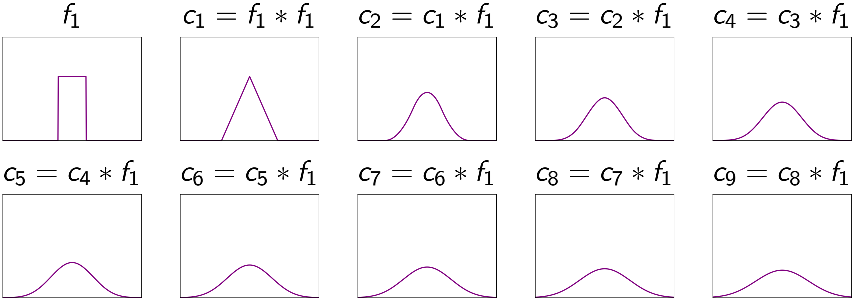
$$\begin{aligned} p_{\tilde{z}}(z) &= \mathrm{P}(\tilde{z} = z) \\ &= \sum_{j=0}^{M-1} \mathrm{P}(\tilde{x} = j, \tilde{y} = z - j) \\ &= \sum_{j=0}^{M-1} \mathrm{P}(\tilde{x} = j)\mathrm{P}(\tilde{y} = z - j) \\ &= \sum_{j=0}^{M-1} p_{\tilde{x}}(j) p_{\tilde{y}}(z - j) \end{aligned}$$

$$p_{\tilde{z}} = p_{\tilde{x}} * p_{\tilde{y}}$$

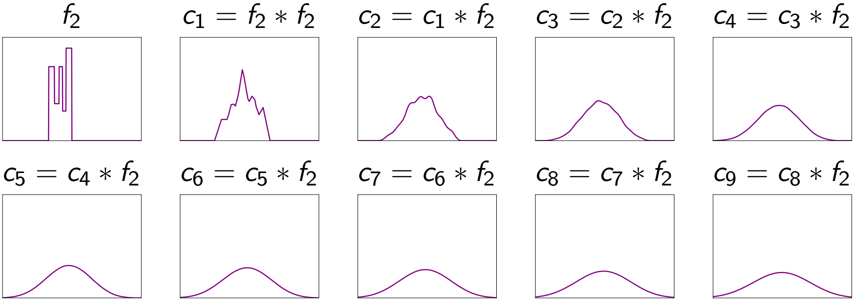
Distribution of sums

Random quantities are often the result of adding many independent components

Iterated convolutions



Iterated convolutions



Central limit theorem

Random quantities are often the result of adding many independent components

Such quantities are Gaussian because iterated convolutions converge to Gaussian functions