DS 4400

Machine Learning and Data Mining I
Spring 2024

David M. Liu
Khoury College of Computer Science
Northeastern University

Tuesday January 16 2024

Today's Outline

- Probability review
  - Random variables (discrete)
  - Expectation and variance
  - Conditional probabilities and independence
  - Bayes Theorem
  - Marginalization
- **Linear algebra review**
  - Matrices
  - Vectors
  - Linear independence
  - Rank of a matrix and matrix inverse

Announcements

- Office hours schedule starting this week.
  - Please create a Khoury user account if you don't have one.
  - For this week only, we will use Zoom (link on Piazza)

Office Hours:

- **Jai:** Tuesdays 2:00 – 3:30pm
- **Dhanush:** Wednesdays 12:00 – 1:30pm
- **David:** Thursdays 2:00 – 3:30pm
- **Caleb:** Fridays 4:00 – 5:30pm

Office hours held virtually on the Khoury Office Hours Application. If you do not have a Khoury account please apply for one here.

- Homework 1 posted to Canvas. Due in 2 weeks on Gradescope. Two problems require coding.

- Please submit project teams by Friday January 26. There will be a forum on Piazza to help with searching for teammates.

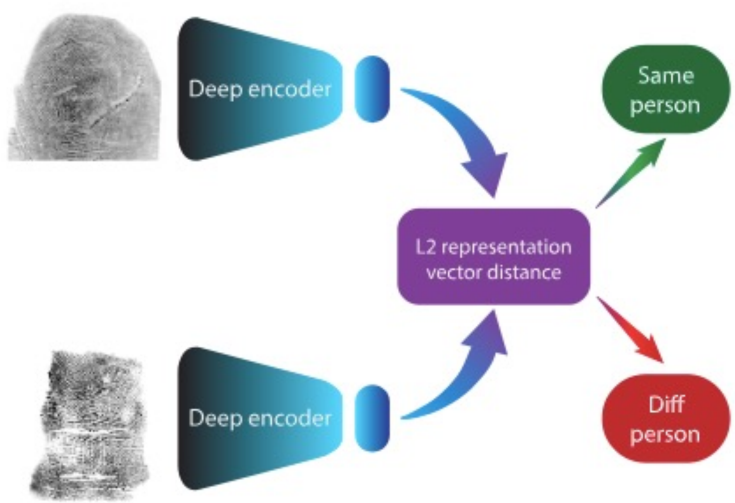
COMPUTER SCIENCE

Unveiling intra-person fingerprint similarity via deep contrastive learning

Gabe Guo[1]*, Aniv Ray[1], Miles Izydorczak[2], Judah Goldfeder[1], Hod Lipson[3], Wenyao Xu[4]

Fingerprint biometrics are integral to digital authentication and forensic science. However, they are based on the unproven assumption that no two fingerprints, even from different fingers of the same person, are alike. This renders them useless in scenarios where the presented fingerprints are from different fingers than those on record. Contrary to this prevailing assumption, we show above 99.99% confidence that fingerprints from different fingers of the same person share very strong similarities. Using deep twin neural networks to extract fingerprint representation vectors, we find that these similarities hold across all pairs of fingers within the same person, even when controlling for spurious factors like sensor modality. We also find evidence that ridge orientation, especially near the fingerprint center, explains a substantial part of this similarity, whereas minutiae used in traditional methods are almost nonpredictive. Our experiments suggest that, in some situations, this relationship can increase forensic investigation efficiency by almost two orders of magnitude.

A



https://www-science-org.ezproxy.neu.edu/content/article/do-prints-two-different-fingers-belong-same-person-ai-can-tell

Probability review

Marginalization

- We can also show that:

$$P(B) = P(B \wedge [A = v_1 \vee A = v_2 \vee \ldots \vee A = v_k])$$

$$\boxed{P(B) = \sum_{i=1}^{k} P(B \wedge A = v_i)} = \sum_{i=1}^{k} P(B \mid A = v_i) P(A = v_i)$$

- This is called **marginalization** over A

EXAMPLE

Marginalization

Consider this *joint* distribution P(A *and* B)

		B		
		1	2	3
A	1	**0.1**	**0.05**	**0.2**
	2	**0.15**	**0**	**0.1**
	3	**0**	**0.2**	**0.2**

$$P(A=2) = P(A=2 \cap B=1) + P(A=2 \cap B=2) + P(A=2 \cap B=3)$$

$$= \sum_{i=1}^{3} P(A=2 \cap B=i)$$

= 0.15 + 0 + 0.1

= 0.25

Recap of Probability

- Discrete Random Variables
  - e.g. Bernoulli random variables
- Expectation and Variance

$$E[X] = \sum_v vPr[X = v] \qquad \mathrm{Var}[X] = E(X^2) - E^2(X)$$

- Conditional Probability and Bayes Rule

$$P(A \mid B) = \frac{P(B \mid A) \times P(A)}{P(B)}$$

- Marginalization

Probability Resources

- Review notes from Stanford's machine learning class
  - http://cs229.stanford.edu/section/cs229-prob.pdf
- David Blei’s probability review
  - https://khoury.neu.edu/home/eelhami/courses/CS6140_Fall16/lecture0_review_probability_1.pdf
- Books:
  - Sheldon Ross, A First course in probability

Linear algebra review

Resources

- Zico Kolter, Linear algebra review
  - http://cs229.stanford.edu/section/cs229-linalg.pdf
- Books:
  - O. Bretscher, Linear Algebra with Applications

Vectors and matrices

- Vector in R^n is an ordered set of n real numbers.
  – e.g. v = (1,6,3,4) is in R^4
  – A column vector: $\begin{pmatrix} 1 \\ 6 \\ 3 \\ 4 \end{pmatrix}$
  – A row vector: $\begin{pmatrix} 1 & 6 & 3 & 4 \end{pmatrix}$
- m-by-n matrix is an object in R^{mxn} with m rows and n columns, each entry filled with a (typically) real number:

$$\begin{pmatrix} 1 & 2 & 8 \\ 4 & 78 & 6 \\ 9 & 3 & 2 \end{pmatrix}$$

Vector operations

- Addition component by component

$$[a_1, a_2, \ldots, a_n] + [b_1, b_2, \ldots, b_n] = [a_1 + b_1, \ldots, a_n + b_n]$$

$$[1, -2,5] + [0,3,7] =$$

- Subtraction is also done component by component

$$[a_1, a_2, \ldots, a_n] - [b_1, b_2, \ldots, b_n] = [a_1 - b_1, \ldots, a_n - b_n]$$

  - Can add and subtract row or column vectors of same dimension
- Dot product
  - Only works for row and column vector of same size

$$[a_1, a_2, \ldots, a_n] \cdot \begin{bmatrix} b_1 \\ \ldots \\ b_n \end{bmatrix} = [a_1 b_1 + \cdots + a_n b_n]$$

$$[1, -2,5] \cdot \begin{bmatrix} 0 \\ 3 \\ 7 \end{bmatrix} =$$

Matrix Operations

Matrix multiplication

We will use upper case letters for matrices. The elements are referred by $A_{i,j}$.

- Matrix product:

$$A \in \mathbb{R}^{m \times n} \qquad B \in \mathbb{R}^{n \times p}$$

$$C = AB \in \mathbb{R}^{m \times p}$$

$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

e.g.

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, B = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

$$AB = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}$$

Matrix transpose

Transpose: You can think of it as

- "flipping" the rows and columns

OR

- "reflecting" vector/matrix on line

e.g. $$\begin{pmatrix} a \\ b \end{pmatrix}^T = \begin{pmatrix} a & b \end{pmatrix}$$

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^T = \begin{pmatrix} a & c \\ b & d \end{pmatrix}$$

A is a symmetric matrix if $A = A^T$

Linear Independence

Linear independence

- A set of vectors is linearly independent if none of them can be written as a *linear combination* of the others.
  - Vectors $x_1,\ldots,x_k$ are linearly independent if $c_1x_1+\ldots+c_kx_k = 0$ implies $c_1=\ldots=c_k=0$
- Otherwise they are linearly dependent

$$x_1 = \begin{pmatrix}1\\2\\1\end{pmatrix} \qquad x_2 = \begin{pmatrix}0\\3\\3\end{pmatrix}$$

Linear independence

- A set of vectors is linearly independent if none of them can be written as a linear combination of the others.
- Vectors $x_1,\ldots,x_k$ are linearly independent if $c_1x_1+\ldots+c_kx_k = 0$ implies $c_1=\ldots=c_k=0$
- Otherwise they are linearly dependent

$$x_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \quad x_2 = \begin{bmatrix} 4 \\ 1 \\ 5 \end{bmatrix} \quad x_3 = \begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$$

Whiteboard for Visual Geometric Understanding

Rank of a Matrix

- rank(A) (the rank of a m-by-n matrix A) is
  - number of linearly independent columns
  - number of linearly independent rows

- If A is n by m, then
  - rank(A)<= min(m,n)

- Examples

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 \\ 4 & 2 \end{pmatrix} \quad \begin{pmatrix} 2 & 1 & 3 \\ 0 & 5 & 2 \end{pmatrix}$$

Matrix Inverse
and
System of Equations

Inverse of a matrix

- Inverse of a square matrix A, denoted by A^{-1} is the *unique* matrix s.t.
  – $AA^{-1} = A^{-1}A = I$ (identity matrix)

- Inverse of a square matrix exists only if the matrix is full rank

- If A^{-1} and B^{-1} exist, then
  – $(AB)^{-1} = B^{-1}A^{-1}$
  – $(A^T)^{-1} = (A^{-1})^T$

Diagonal matrices

A diagonal matrix is a square matrix with zeros everywhere except the diagonal.

$$D = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

$D^T = D$

D^{-1} : replace all entries of the diagonal with their reciprocal

$$D^{-1} = \begin{bmatrix} \frac{1}{3} & 0 & 0 \\ 0 & \frac{1}{9} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix}$$

System of linear equations

$$\begin{array}{rcrcr} 4x_1 & - & 5x_2 & = & -13 \\ -2x_1 & + & 3x_2 & = & 9. \end{array}$$

Matrix formulation

$$Ax = b$$

$$A = \left[\begin{array}{cc} 4 & -5 \\ -2 & 3 \end{array} \right], \quad b = \left[\begin{array}{c} -13 \\ 9 \end{array} \right].$$

If A has an inverse, solution is $x = A^{-1}b$

Recap of Linear Algebra

- Vector and Matrix Operations
  - Add, multiply, transpose
- Linear Independence
- Rank
- Matrix Inverse and System of Equations
- Continuous Random Variables
  - PDFs and CDFs

Not covered in this lecture: determinants, eigenvalues/eigenvectors, matrix calculus

For more, see Stanford CS 229 notes pages 1-12

Announcements

- Office hours schedule starting this week.
  - Please create a Khoury user account if you don't have one.
  - For this week only, we will use Zoom (link on Piazza)

Office Hours:

- **Jai:** Tuesdays 2:00 – 3:30pm
- **Dhanush:** Wednesdays 12:00 – 1:30pm
- **David:** Thursdays 2:00 – 3:30pm
- **Caleb:** Fridays 4:00 – 5:30pm

Office hours held virtually on the Khoury Office Hours Application. If you do not have a Khoury account please apply for one here.

- Homework 1 posted to Canvas. Due in 2 weeks on Gradescope. Two problems require coding.

- Please submit project teams by Friday January 26. There will be a forum on Piazza to help with searching for teammates.