NC STATE UNIVERSITY

CSC-591: Foundations of Data Science
T/Th. 12:50-2:05pm. EBI-1005.

Ranga Raju Vatsavai
Chancellors Faculty Excellence Associate Professor in Geospatial Analytics
Department of Computer Science, North Carolina State University (NCSU)
Associate Director, Center for Geospatial Analytics, NCSU
&
Joint Faculty, Oak Ridge National Laboratory (ORNL)

W3: 9/1-3/15

NC STATE UNIVERSITY

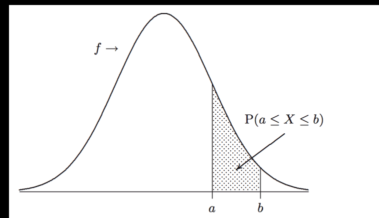
Administrative

- Books and Lecture Notes
- Introductory Statistics
  - https://www.openintro.org/stat/
- Linear and Matrix Algebra
  - http://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf
  - http://www.mathworks.com/moler/eigs.pdf

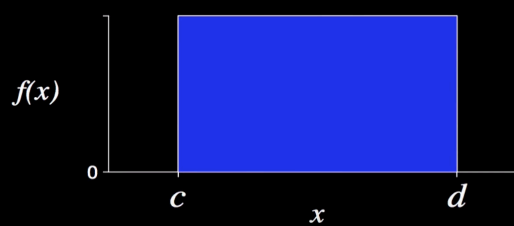
NC STATE UNIVERSITY

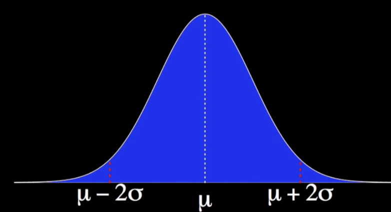
Key points from 9/1

- Continuous distributions
- Uniform: U(c,d)
- Normal: N(μ, σ^2)



Area under a probability density function f on the interval $[a, b]$.





NC STATE UNIVERSITY

Today

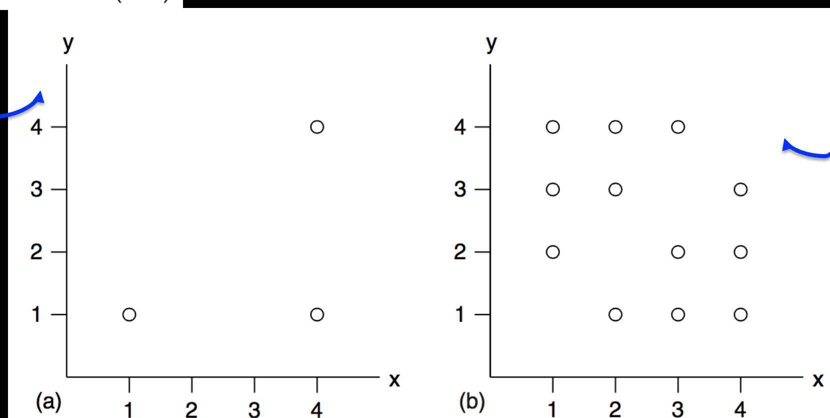
- Basic Linear Algebra and Calculus

Why?

- Let us consider the following datasets

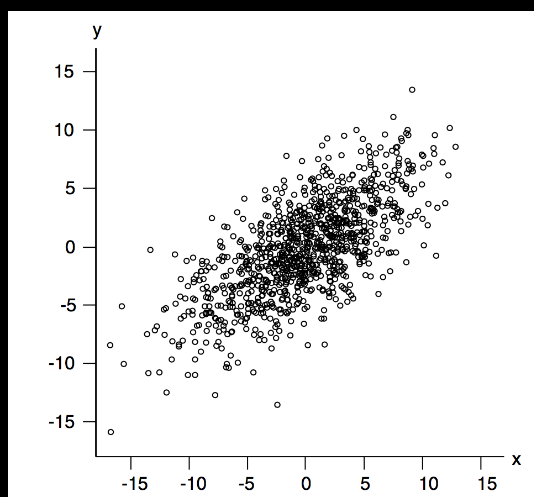
$$\mathbf{A} = (\mathbf{x} : \mathbf{y}) = \begin{pmatrix} 1 & 1 \\ 4 & 1 \\ 4 & 4 \end{pmatrix}$$

$$\mathbf{B}' = (\mathbf{x} : \mathbf{y})' = \begin{pmatrix} \mathbf{x}' \\ \mathbf{y}' \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 & 3 & 4 & 4 & 4 \\ 2 & 3 & 4 & 1 & 3 & 4 & 1 & 2 & 4 & 1 & 2 & 3 \end{pmatrix}$$



Why

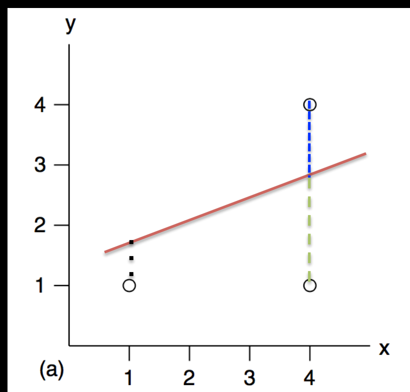
- Let us consider the following datasets

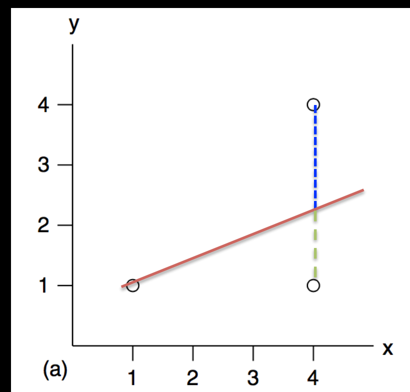


1000 observations from $\mathrm{N}_2(\mathbf{0}, \boldsymbol{\Sigma})$; $\sigma_x = 5$, $\sigma_y = 4$, $\varrho_{xy} = 0.7$.

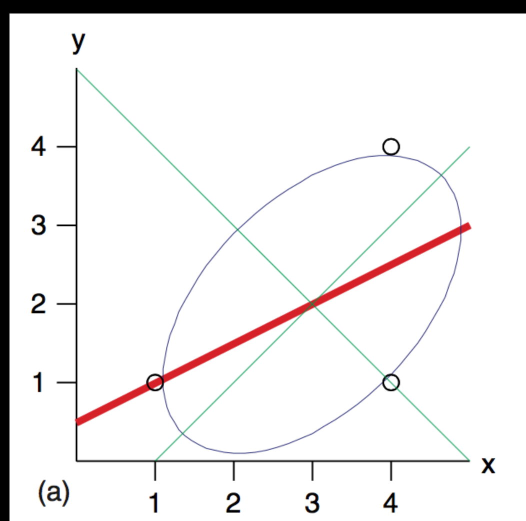
Question

- Draw a line $\hat{y} = \beta x + \varepsilon$ into each scatter plot such that the sum of squared vertical distances $(y_i - \hat{y}_i)^2$ would be as small as possible





Solution



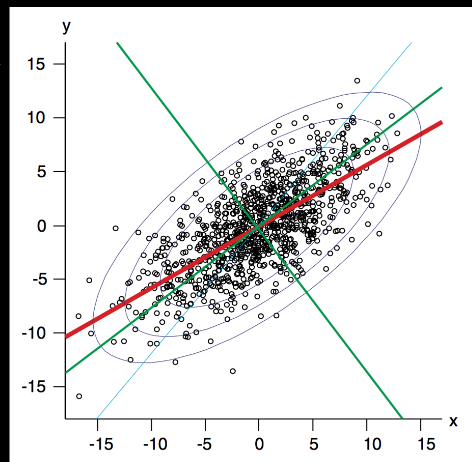
NC STATE UNIVERSITY

Solution

Assume Normal Distribution:
$N_2(\mathbf{0}, \mathbf{\Sigma})$; $\sigma_x = 5$, $\sigma_y = 4$, $\rho_{xy} = 0.7$.

Regression line has the slope $\beta \approx \rho_{xy}\sigma_y/\sigma_x$

Direction of first major axis of contour ellipse is determined by the first eigenvector of $\mathbf{\Sigma}$



9/8/15 © Raju Vatsavai CSC-591. 9

NC STATE UNIVERSITY

Linear Algebra

- Linear algebra is the study of linear maps on finite-dimensional vector spaces.

9/8/15 © Raju Vatsavai CSC-591. 10

NC STATE UNIVERSITY

Complex Numbers

Definition ***complex numbers***

- A ***complex number*** is an ordered pair (a, b), where $a, b \in \mathbf{R}$, but we will write this as $a + bi$.
- The set of all complex numbers is denoted by $\mathbf{C}$:

$$\mathbf{C} = \{a + bi : a, b \in \mathbf{R}\}.$$

- ***Addition and multiplication*** on $\mathbf{C}$ are defined by

$$(a + bi) + (c + di) = (a + c) + (b + d)i,$$
$$(a + bi)(c + di) = (ac - bd) + (ad + bc)i;$$

here $a, b, c, d \in \mathbf{R}$.

NC STATE UNIVERSITY

Example

- (2+3i)(4+5i)
- -7 + 22i

NC STATE UNIVERSITY

Properties of Complex Arithmetic

commutativity
$\alpha + \beta = \beta + \alpha$ and $\alpha\beta = \beta\alpha$ for all $\alpha, \beta \in \mathbf{C}$;

associativity
$(\alpha + \beta) + \lambda = \alpha + (\beta + \lambda)$ and $(\alpha\beta)\lambda = \alpha(\beta\lambda)$ for all $\alpha, \beta, \lambda \in \mathbf{C}$;

identities
$\lambda + 0 = \lambda$ and $\lambda 1 = \lambda$ for all $\lambda \in \mathbf{C}$;

additive inverse
for every $\alpha \in \mathbf{C}$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha + \beta = 0$;

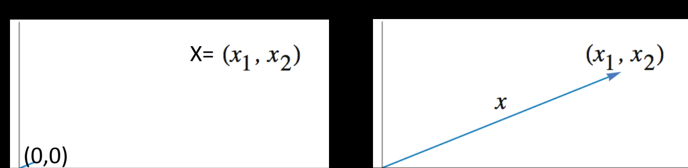
multiplicative inverse
for every $\alpha \in \mathbf{C}$ with $\alpha \neq 0$, there exists a unique $\beta \in \mathbf{C}$ such that $\alpha\beta = 1$;

distributive property
$\lambda(\alpha + \beta) = \lambda\alpha + \lambda\beta$ for all $\lambda, \alpha, \beta \in \mathbf{C}$.

NC STATE UNIVERSITY

Vector

- [1.2, 2.5]
  - Is a 2-vector over R, written as R^2
  - General notation, R^d
- A vector can be thought as a function
- F^d : {0,1,2,…,d-1} -> F
- Sparse vector
  - If most elements are 0's



Elements of R^2 can be thought of as points or vectors

NC STATE UNIVERSITY

What kind of data

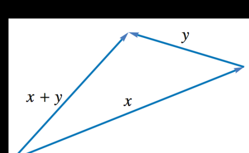
- Document as bag of words
  - f : WORDS -> R
- Collections of attributes
  - Physical characteristics of persons
  - Demographic records of customers
- Probability distribution, e.g., {1:1/6, 2:1/6, …, 6:1/6}
- Images
  - {r, g, b}

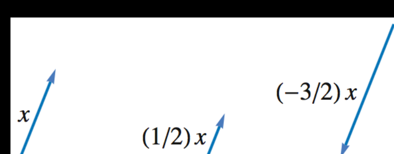
9/8/15 © Raju Vatsavai CSC-591. 15

NC STATE UNIVERSITY

Vector Operations

- Complex numbers, translation is achieved by adding a complex number. f(z) = z + (1+2i)
- Vector addition: add element-wise $[u_1,u_2] + [v_1,v_2] = [u_1+v_1, u_2+v_2]$
- Scalar-vector multiplication (is a scaling operation): αv





9/8/15 © Raju Vatsavai CSC-591. 16

NC STATE UNIVERSITY

Convex Combinations

- For given $0 \le \alpha \le 1$, and $0 \le \beta \le 1$, and $\alpha+\beta=1$; an expression of form: $\alpha u + \beta v$ is called convex combination of u and v
- Application (average of two images)
- Lets say u and v are two images, and $\alpha=0.5$ and $\beta=0.5$, the $\alpha u + \beta v$ is average of two images

NC STATE UNIVERSITY

Vector Operations

- Dot product of two D-vectors is sum of the product of corresponding entries
  - $u.v = \Sigma_{k \epsilon D} u[k]v[k]$
- What do you think the output is?
- Examples
  - Total cost of a product
  - Measuring similarity

NC STATE UNIVERSITY

Vector Space

A ***vector space*** is a set V along with an addition on V and a scalar multiplication on V such that the following properties hold:

commutativity
$u + v = v + u$ for all $u, v \in V$;

associativity
$(u + v) + w = u + (v + w)$ and $(ab)v = a(bv)$ for all $u, v, w \in V$ and all $a, b \in \mathbf{F}$;

additive identity
there exists an element $0 \in V$ such that $v + 0 = v$ for all $v \in V$;

additive inverse
for every $v \in V$, there exists $w \in V$ such that $v + w = 0$;

multiplicative identity
$1v = v$ for all $v \in V$;

distributive properties
$a(u + v) = au + av$ and $(a + b)v = av + bv$ for all $a, b \in \mathbf{F}$ and all $u, v \in V$.

NC STATE UNIVERSITY

Vector Space F^S

- If S is a set, then $\mathbf{F}^S$ denotes the set of functions from S to $\mathbf{F}$.
- For $f, g \in \mathbf{F}^S$, the ***sum*** $f + g \in \mathbf{F}^S$ is the function defined by
$$(f + g)(x) = f(x) + g(x)$$
for all $x \in S$.
- For $\lambda \in \mathbf{F}$ and $f \in \mathbf{F}^S$, the ***product*** $\lambda f \in \mathbf{F}^S$ is the function defined by
$$(\lambda f)(x) = \lambda f(x)$$
for all $x \in S$.
- F^S is a vector space
- *In general, a vector space is an abstract entity whose elements might be lists, functions, or weird objects.*

NC STATE UNIVERSITY

Subspace

- A subset U of V is called a *subspace* of V if U is also a vector space (using the same addition and scalar multiplication as on V).
- A subset U of V is a subspace of V if and only if U satisfies the following three conditions:
  - Additive identity: 0 ε U
  - Closed under addition: u, w ε U => u + w ε U
  - Closed under scalar multiplication

    a ε F, u ε U => au ε U

9/8/15 © Raju Vatsavai CSC-591. 21

NC STATE UNIVERSITY

Important properties of subspaces

Sum of subspaces is the smallest containing subspace

Suppose $U_1, \dots, U_m$ are subspaces of V. Then $U_1 + \cdots + U_m$ is the smallest subspace of V containing $U_1, \dots, U_m$.

Direct Sum of Subspaces

Suppose $U_1, \dots, U_m$ are subspaces of V.

- The sum $U_1 + \cdots + U_m$ is called a ***direct sum*** if each element of $U_1 + \cdots + U_m$ can be written in only one way as a sum $u_1 + \cdots + u_m$, where each u_j is in U_j.
- If $U_1 + \cdots + U_m$ is a direct sum, then $U_1 \oplus \cdots \oplus U_m$ denotes $U_1 + \cdots + U_m$, with the $\oplus$ notation serving as an indication that this is a direct sum.

9/8/15 © Raju Vatsavai CSC-591. 22

NC STATE UNIVERSITY

Linear combination of vectors

- An expression: $\alpha_1 v_1 + \ldots + \alpha_n v_n$
is a linear combination of vectors $v_1 \ldots v_n$ and scalars $\alpha_1 \ldots \alpha_n$ are coefficients of the linear combinations
- Examples
  - Given a set of raw materials, different products can be made with different combinations of materials, then total resource utilization is linear combination of materials for each product ($u = \alpha_1 v_1 + \ldots + \alpha_n v_n$)
  - In F^3 (17,4,2) is a linear combination of (2,1,-3), (1,-2,4) because
    - (17, -4, 2) = 6(2,1,-3) + 5(1,-2,4)

9/8/15 © Raju Vatsavai CSC-591. 23

NC STATE UNIVERSITY

Span

The set of all linear combinations of a list of vectors $v_1, \ldots, v_m$ in V is called the ***span*** of $v_1, \ldots, v_m$, denoted $\text{span}(v_1, \ldots, v_m)$. In other words,

$$\text{span}(v_1, \ldots, v_m) = \{a_1 v_1 + \cdots + a_m v_m : a_1, \ldots, a_m \in \mathbf{F}\}.$$

The span of the empty list ( ) is defined to be $\{0\}$.

- Previous example shows that in F3
  - (17, -4, 2) ε span((2,1,-3), (1,-2,4))
- Spans
  - If span (v_1, …v_m) equals V, we say that v_1, …, v_m spans V

9/8/15 © Raju Vatsavai CSC-591. 24

NC STATE UNIVERSITY

Linearly independent

- A list $v_1, \ldots, v_m$ of vectors in V is called ***linearly independent*** if the only choice of $a_1, \ldots, a_m \in \mathbf{F}$ that makes $a_1 v_1 + \cdots + a_m v_m$ equal 0 is $a_1 = \cdots = a_m = 0$.
- The empty list ( ) is also declared to be linearly independent.

NC STATE UNIVERSITY

Bases

- A basis of V is a list of vectors in V that is linearly independent and spans V.
- Examples
  - The list (1,0, …,0),(0,1,…,0), …, (0,…,0,1) is a basis of F^n, called the standard basis of F^n
  - The list (1,2),(3,5) is a basis of F^2

Criteria for basis

- A list $v_1, \ldots, v_n$ of vectors in V is a basis of V if and only if for every $v \in V$ can be written uniquely in the form $v = a_1v_1 + \ldots + a_nv_n$, where $a_1, \ldots, a_n \in F$.

- Every spanning list in a vector space can be reduced to basis of the vector space
- Every finite-dimensional vector space has a basis

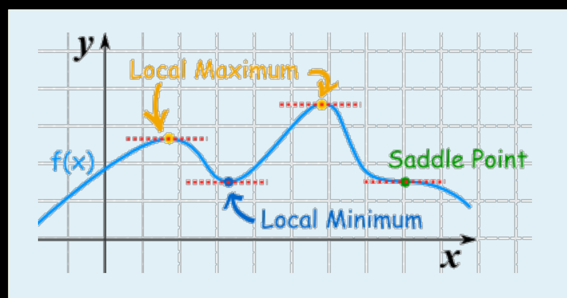
Matrix Algebra

- Addition
- Multiplication
- Inverse

NC STATE UNIVERSITY

Basic Calculus

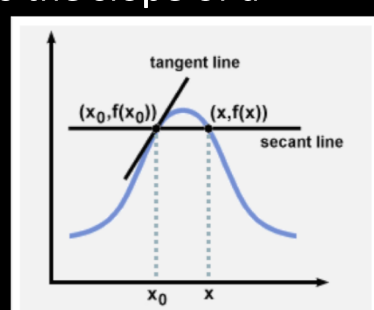
- How to finding maximum and minim of a function



NC STATE UNIVERSITY

Derivative

- The derivative of a function f (x) at x = x_0 , denoted f'(x_0) or (x_0) , can be naively defined as the slope of the graph of f at x = x_0.
- Derivative essentially finds the slope of a function



NC STATE UNIVERSITY

Acknowledgements

- Vipin Kumar (Minnesota)
- Jiawei Han (UIUC)
- Hanspeter Pfister (Harvard)
- Larry Wasserman (CMU)

