Elements of Computational Biology

Saul Pierotti

October 23, 2019

21/10/19

- Linear algebra is one of the main topics for developping algorithms
- Biology ha an high level of noise in its data
- I should know the formula of binomial and normal distribution
- Tests like T-Student, ANOVA
- Slides: ww.biocomp.unibo.it/gigi/2019-2020/ECB

Vectors

- Given a reference system, a vector is represented by its components on the axis
- $\vec{x} \in R^n$ means x is a real vector in an n-dimensional space
- $\vec{x} \in C^n$ means x is a complex vector in an n-dimensional space
- Sum of vectors is done by summing their components or graphically with the parallelogram rule
  - $\vec{c} = \vec{a} + \vec{b} means c_i = a_i + b_i for i = 1..n$
  - Difference is the same concept
  - $For every \vec{v} there is the \vec{0} for which the sum \vec{v} + \vec{0} = \vec{v}$
  - You can only sum vectors in the same vector space
- Scalar multiplication
  - $\vec{c} = \lambda \vec{a} means c_i = a_i \lambda$
  - A scalar multiplication of a sum is the sum of the scalar multiplications of the components
- The norm of a vector $||\vec{v}||$ is its lenght
  - Can be computed with the pitagorean theorem $||\vec{v}|| = \sqrt{\sum i = 1nv_1^2}$
  - The norm of the sum is less or equal to the sum of the norm of the components
  - The scalar product of a norm is the norm of the scalar product
  - Norm in higher dimensions
    - Pitagorean theorem with all the components
- Distance between points in space is the norm of the difference between the vectors defining the points
- Dot product, also called scalar or inner product
  - You can use the notation $< A, B >$
  - It is used in physics to calculate work
  - $\vec{w} = ||\vec{F}||||\vec{s}||cos\theta = \sum i = 1nF_i * s_i$
  - It is a number, complex or real depending on the vectors!
  - It is commutative and distributive
  - $< x, x >= ||\vec{x}||^2$
  - It is positive when the angle is acute
  - No cancelation rule
    - $< A, B >=< A, C > does NOT imply \vec{B} = \vec{C}$
- Angle between vectors
  - Can be calculated inverting the dot product
- Line in passing through the origin

- Can be defined as the set of points ortoghonal to a vector $\vec{w}$
- $w_1x_1 + w_2x_2 = 0$
- In higher dimensions this describes an hyperplane (an n-1 dimensional object)

22/10/19

- All the point on an hyperplane have the same projection on its defining vector $\vec{w}$
- The projection of $\vec{b}$ on $\vec{a}$ is calculated as $\vec{b} * cos\theta$
- $< \vec{w}, y\vec{x} > +b = 0$ where $b = -\frac{p}{\vec{w}}$. with p being the projection
- A projection is a number which is positive when $\vec{w}$
- 2 hyperplanes are parallel if their are defined by the same vector $\vec{w}$ allowing for a scaling factor λ
- The distance between parallel hyperplanes is computed as the difference of their projections on $\vec{w}$
- The distance of a point A from an hyperplane is the projection of the point on the defining vector $\vec{w}$, minus the projection of the hyperplane on the same vector
  - This is $\frac{<\vec{A}, \vec{w}>}{||\vec{w}||}$
- Hyperplanes are useful for the separation of classes of data # Matrices
- A matrix is an array of numbers arranged in a rectangural structure
- It has n rows and m columns, it is represented as $A \in R^{n*m}$
- The index is always nm, meaning row and then column
- We can sum only matrices of the same dimensions, they are said to be conformable for addition
- The sum is defined as the sum of the respective elements
- The 0 matrix contains all 0 elements and does not change the matrix it is added to
- Scalar multiplication is performed multiplying all the elements of the matrix for the scalar
  - $C = \lambda A$ implies $c_{ij} = \lambda a_{ij}$
- Matrix addition and scalar multiplication are commutative, associative and distributive
- Matrix product has peculiar properties
  - $C = A * B$ can be computed as row by column product
  - It can be defined only if the number of columns in the first matrix is equal to the number of rows of the second
  - The result is a matrix with the same number of rows as the first, and the same number of columns as the second
  - Do exercised because they are so important
  - The product between matrices is NOT commutative!
  - $A(B + C) = AB + AC$
  - $(A + B)C = AC + BC$
  - $A(BC) = (AB)C$
  - Be aware!
    * If $AB = 0$ we can NOT conclude that B or C are 0
    * If $AB = AC$ we can NOT conclude that $B = C$
- An upper triangular matrix has all the elements below the diagonal equal to 0, and a lower triangular the ones above it
- A diagonal matrix has all the elements outside the diagonal equal to 0
- A diagonal matrix with all 1 elements is the identity matrix I
  - It does not change the square matrix it is multiplied to
  - In this case, $AI = IA = A$
- In numbers, an invers of a is the number b such that $ab = 1$, so $b = 1/a$
- The inverse of a matrix, called A^{-1}, can exist but that is not always the case
- The transposition of a nm *matrix is a* mn matrix, called A^t, where $[A^t]_{ij} = A_{ji}$
  - A and A^t are always conformable to product, in both directions
- A matrix is symmetric if $A = A^t$, antisymmetric if $A = -A^t$
- An antisymmetric matrix has a 0 diagonal and antisymmetrical elements otherwise
- An orthogonal matrix has its invers equal to the transposal, $A^{-1} = A^t$
  - They are really useful because they can describe a spatial rotation
  - You can check for orthoganlity by checking that $A * A^t = I$

- Some properties of transpose and inverse matrices
  - $(AB)^{-1} = B^{-1} * A^{-1}$, but only if $(AB)^{-1}$ exists(!)
  - $(AB)^t = B^t * A^t$
- It is possible to associate a number called determinant to any square matrix
  - $det(A) \in R$
  - For an order 2 square matrix, that is $det(A) = a_{11} * a_{22} - a_{12} * a_{21}$

23/10/19

- If an entire row or column is equal to 0, then the determinant of the matrix is 0
- $det(A * B) = det(A) * det(B)$
- The determinant of an orthogonal matrix is either 1 or -1
- For 2*2 matrices
  - $A * A^{-1} = I$ if $det(A) \neq 0$
  - $A^{-1} = \frac{1}{det(A)} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$
- A minor M_{ij} of a matrix is the determinant of any square submatrix of A
- The cofactor of the element a_{ij} is $C_{ij} = M_{ij} * (-1)^{i+j}$
- To compute the determinant of any matrix you pick any row or column and sum the product of any element in it for its cofactor
  - In a column $det(A) = \sum\limits_{i=1}^{n} a_{ij} * C_{ij}$
  - In a row $det(A) = \sum\limits_{j=1}^{n} a_{ij} * C_{ij}$
  - It is convenient to choose the row or column with most 0 for the computation
  - If 2 rows are identical, det(A)=0
  - If one row is 0, then det(A)=0
  - If you exchange 2 rows, det(A')=-det(A)
  - The determinant of a triangular matrix si the product of the diagonal elements
- Matrix can represent systems of linear equations
  - The system $\begin{cases} x + y = 7 \\ 3x - y = 5 \end{cases}$ can be represented as $\begin{pmatrix} 1 & 1 \\ 3 & -1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 7 \\ 5 \end{pmatrix}$
  - The system has a solution if a matrix is invertible
- Every column of a matrix can be thought of as a vector
  - To make the dot product of 2 vectors using matrices you can multiply one vector for the traspose of the second
  - $< \vec{a}, \vec{b} >= A * B^t$
- A matrix can be thought as a linear transformation of a vector
  - $A^{3*2} : R^2 \to R^3$