# Distributional Semantics

Raffaella Bernardi

University of Trento

March, 2017

# Linear equation

**A linear equation** is an algebraic equation in which each term is either a constant or the *product of a constant and a single* variable. E.g. a two variables $x$ and $y$ is $y = mx + b$, where $m$ and $b$ designate constants.
The origin of the name "*linear*" comes from the fact that the set of solutions of such an equation forms a *straight line* in the plane.
The general linear equation in n variables is:
$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$
In this form, $a_1, a_2, \ldots, a_n$ are the coefficients, $x_1, x_2, \ldots, x_n$ are the variables, and $b$ is the constant.
Such an equation will represent an $(n-1)$-dimensional hyperplane in $n$-dimensional Euclidean space (or in our case n-dimensional vector space)

# Background: Vector and Matrix

## Vector Space

**A vector space** is a mathematical structure formed by a collection of vectors: objects that may be added together and multiplied (“scaled”) by numbers, called scalars in this context.

**Vector** an n-dimensional vector is represented by a column:

$$\begin{bmatrix} v_1 \\ \dots \\ v_n \end{bmatrix}$$

or for short as $\vec{v} = (v_1, \dots v_n)$.

# Background: Vector and Matrix

## Operations on vectors

Vector addition:

$$\vec{v} + \vec{w} = (v_1 + w_1, \ldots v_n + w_n)$$

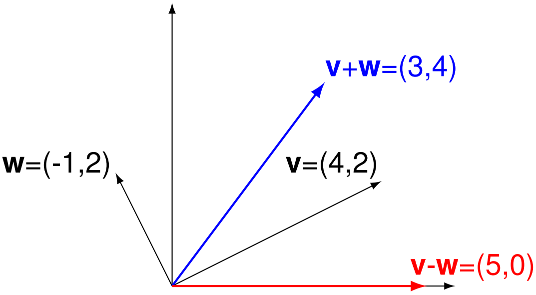
similarly for the $-$.

Scalar multiplication: $c\vec{v} = (cv_1, \ldots cv_n)$ where $c$ is a “scalar”.

# Background: Vector and Matrix

## Vector visualization

Vectors are visualized by arrows. They correspond to points (the point where the arrow ends.)




vector addition produces the diagonal of a parallelogram.

# Background: Vector

## Length

**Length** $||\vec{v}|| = \sqrt{\vec{v} \cdot \vec{v}} = \sqrt{\sum_{i=1}^{n} v_i^2}$

| | | d1, | d2, | d3, | d4, | d5, | d6, | d7, | d8 |
|---|---|---|---|---|---|---|---|---|---|
| w1 | = | (3, | 3, | 3, | 3, | 0, | 0, | 0, | 0) |
| w2 | = | (1, | 1, | 1, | 1, | 0, | 0, | 0, | 0) |
| w3 | = | (0, | 0, | 0, | 0, | 1, | 1, | 1, | 1) |

$$||\overrightarrow{w1}|| = \sqrt{9+9+9+9} = \sqrt{36} = 6$$
$$||\overrightarrow{w2}|| = \sqrt{1+1+1+1} = \sqrt{4} = 2$$
$$||\overrightarrow{w3}|| = \sqrt{1+1+1+1} = \sqrt{4} = 2$$

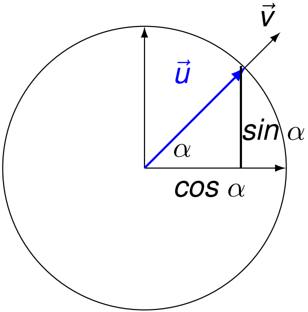
# Background: Vector

## Unit vector

**Unit vector** is a vector whose length equals one.

$$\vec{u} = \frac{\vec{v}}{||\vec{v}||}$$

is a unit vector in the same direction as $\vec{v}$. (normalized vector)




$\vec{u} = \frac{\vec{v}}{||\vec{v}||} = (cos\ \alpha, sin\ \alpha)$

# Background: Vectors

## Dot product or inner product

$$\vec{v} \cdot \vec{w} = (v_1 w_1 + \ldots + v_n w_n) = \sum_{i=1}^{n} v_i w_i$$

**Example** We have three goods to buy and sell, their prices are $(p_1, p_2, p_3)$ (price vector $\vec{p}$). The quantities we are buy or sell are $(q_1, q_2, q_3)$ (quantity vector $\vec{q}$, their values are positive when we sell and negative when we buy.) Selling the quantity $q_1$ at price $p_1$ brings in $q_1 p_1$. The total income is the *dot product*:

$$\vec{q} \cdot \vec{p} = (q_1, q_2, q_3) \cdot (p_1, p_2, p_3) = q_1 p_1 + q_2 p_2 + q_3 p_3$$

# Inner product: example

| | d1, | d2, | d3, | d4, | d5, | d6, | d7, | d8 |
|---|---|---|---|---|---|---|---|---|
| w1 | 3, | 3, | 3, | 3, | 0, | 0, | 0, | 0 |
| w2 | 1, | 1, | 1, | 1, | 0, | 0, | 0, | 0 |

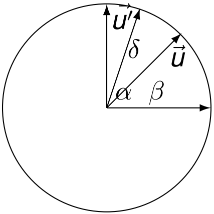
$$\overrightarrow{w1} \cdot \overrightarrow{w2} = \sum_i w1_i \times w2_i = (3 \times 1) + (3 \times 1) + (3 \times 1) + (3 \times 1) = 12$$

# Background: Vector

## Cosine formula

Given $\delta$ the angle formed by the two unit vectors $\vec{u}$ and $\vec{u'}$, s.t.
$\vec{u} = (cos\ \beta, sin\ \beta)$ and $\vec{u'} = (cos\ \alpha, sin\ \alpha)$

$$\vec{u} \cdot \vec{u'} = (cos\ \beta)(cos\ \alpha) + (sin\ \beta)(sin\ \alpha) = cos(\beta - \alpha) = cos\ \delta$$




Given two arbitrary vectors *v* and *w*:

$$cos\ \delta = \frac{\vec{v}}{||\vec{v}||} \cdot \frac{\vec{w}}{||\vec{w}||}$$

The bigger the angle $\delta$, the smaller is $cos\ \delta$; $cos\ \delta$ is never bigger than 1 (since we used unit vectors) and never less than -1. It's 0 when the angle is $90^o$

# Cosine Similarity

## Example

$$cos(\vec{x},\vec{y}) = \frac{\vec{x}\cdot\vec{y}}{||\vec{x}||||\vec{y}||} = \frac{\sum_{i=1}^{n} x_i \times y_i}{\sqrt{\sum_{i=1}^{n} x_i^2} \times \sqrt{\sum_{i=1}^{n} y_i^2}}$$

| | | d1, | d2, | d3, | d4, | d5, | d6, | d7, | d9 |
|---|---|---|---|---|---|---|---|---|---|
| w1 | = | (3, | 3, | 3, | 3, | 0, | 0, | 0, | 0) |
| w2 | = | (1, | 1, | 1, | 1, | 0, | 0, | 0, | 0) |
| w3 | = | (0, | 0, | 0, | 0, | 1, | 1, | 1, | 1) |

$$cos(\vec{w}1,\vec{w}2) = \frac{\vec{w}1\cdot\vec{w}2}{||\vec{w}1||\times||\vec{w}2||} = \frac{(3\times1)+(3\times1)+(3\times1)+(3\times1)}{6\times2} = \frac{12}{12}$$

$$cos(\vec{w}1,\vec{w}3) = \frac{\vec{w}1\cdot\vec{w}3}{||\vec{w}1||\times||\vec{w}3||} = \frac{0}{6\times2} = \frac{0}{12} = 0$$

# Background: Matrix

## Matrices multiplication

A matrix is represented by [nr-rows x nr-columns].
Eg. for a 2 x 3 matrix, the notation is:

$$A = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix}$$

$a_{ij}$ $i$ stands for the row nr, and $j$ stands for the column nr.

The multiplication of two matrices is obtained by

*Rows of the 1st matrix x columns of the 2nd.*

A matrix with m-columns can be multiplied only by a matrix of m-rows:

[n x m] x [m x k ] = [n x k].

# Background: Vector and Matrix

## A matrix acts on a vector

Example of 2 x 2 matrix multiplied by a 2 x 1 matrix (viz. a vector). Take $A$ and $\vec{x}$ to be as below.

$$A\,\vec{x} = \begin{bmatrix} 1 & 0 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} (1,0)\cdot(x_1,x_2) \\ (-1,1)\cdot(x_1,x_2) \end{bmatrix} = \begin{bmatrix} 1(x_1)+0(x_2) \\ -1(x_1)+1(x_2) \end{bmatrix} =$$

$$= \begin{bmatrix} x_1 \\ x_2 - x_1 \end{bmatrix} = \vec{b}$$

$A$ is a “difference matrix”: the output vector $\vec{b}$ contains differences of the input vector $\vec{x}$ on which “the matrix has acted.”