Quick Tour of Linear Algebra and Graph Theory

CS224w: Social and Information Network Analysis
Fall 2012
Peter Lofgren
Based on Yu ”Wayne” Wu and Borja Pelato’s previous versions

Linear Function

A linear function M is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ that satisfies two properties:

1. For all $x, y \in \mathbb{R}$,

$$M(x + y) = M(x) + M(y)$$

2. For all $x \in \mathbb{R}$ and all $a \in \mathbb{R}$

$$M(ax) = aM(x)$$

Linear Function

A linear function M is a function from $\mathbb{R}^n$ to $\mathbb{R}^m$ that satisfies two properties:

1. For all $x, y \in \mathbb{R}$,

$$M(x + y) = M(x) + M(y)$$

2. For all $x \in \mathbb{R}$ and all $a \in \mathbb{R}$

$$M(ax) = aM(x)$$

Every linear function can be represented by a matrix. Every matrix is a linear function.

Matrices and Vectors

- Matrix: A rectangular array of numbers, e.g., $A \in \mathbb{R}^{m \times n}$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

- Vector: A matrix consisting of only one column (default) or one row, e.g., $x \in \mathbb{R}^n$

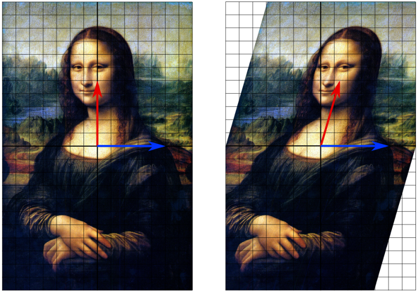
$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Transform Example

Let

$$M = \begin{bmatrix} 1 & 0.3 \\ 0 & 1 \end{bmatrix}.$$

If we apply M to every point on the Mona Lisa, we get the following:



Matrix Multiplication

- If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, then their product $AB \in \mathbb{R}^{m \times p}$ is the unique matrix such that for any $x \in \mathbb{R}^p$,

$$(AB)(x) = A(B(x)).$$

Matrix Multiplication

- If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, then their product $AB \in \mathbb{R}^{m \times p}$ is the unique matrix such that for any $x \in \mathbb{R}^p$,

$$(AB)(x) = A(B(x)).$$

We can compute the product $C = AB$ using this formula:

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

Matrix Multiplication

- If $A \in \mathbb{R}^{m \times n}$, $B \in \mathbb{R}^{n \times p}$, then their product $AB \in \mathbb{R}^{m \times p}$ is the unique matrix such that for any $x \in \mathbb{R}^p$,

$$(AB)(x) = A(B(x)).$$

We can compute the product $C = AB$ using this formula:

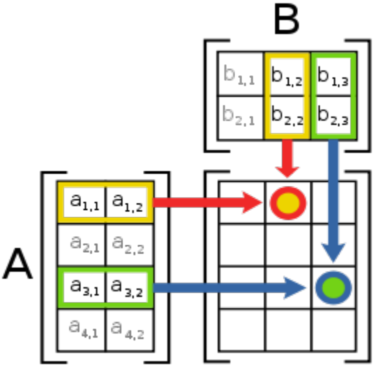
$$C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$$

- Special cases: Matrix-vector product, inner product of two vectors. e.g., with $x, y \in \mathbb{R}^n$:

$$x^T y = \sum_{i=1}^{n} x_i y_i \in \mathbb{R}$$

Matrix Multiplication

$$(AB)(x) = A(B(x))$$



Properties of Matrix Multiplication

- Associative: $(AB)C = A(BC)$
- Distributive: $A(B + C) = AB + AC$
- Non-commutative: $AB \neq BA$

Operators and properties

- Transpose: $A \in \mathbb{R}^{m\times n}$, then $A^T \in \mathbb{R}^{n\times m}$: $(A^T)_{ij} = A_{ji}$.
  For example, if
  $$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$
  then
  $$A^T = \begin{bmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{bmatrix}$$
- Properties:
  - $(A^T)^T = A$
  - $(AB)^T = B^T A^T$
  - $(A+B)^T = A^T + B^T$

Identity Matrix

- Identity matrix: $I = I_n \in \mathbb{R}^{n \times n}$:

$$I_{ij} = \begin{cases} 1 & \text{i=j,} \\ 0 & \text{otherwise.} \end{cases}$$

- $\forall A \in \mathbb{R}^{m \times n}$: $AI_n = I_m A = A$

$$I_1 = \begin{bmatrix} 1 \end{bmatrix},\ I_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},\ I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix},\ \cdots,\ I_n = \begin{bmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}$$

Diagonal Matrix

- Diagonal matrix: $D = diag(d_1, d_2, \ldots, d_n)$:

$$D_{ij} = \begin{cases} d_i & \text{j=i,} \\ 0 & \text{otherwise.} \end{cases}$$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & -3 \end{bmatrix}$$

Other Special Mtrices

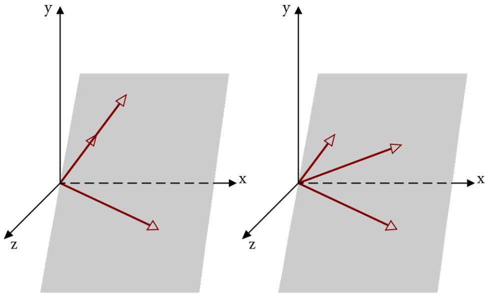
- Symmetric matrices: $A \in \mathbb{R}^{n \times n}$ is symmetric if $A = A^T$.
- Orthogonal matrices: $U \in \mathbb{R}^{n \times n}$ is orthogonal if $UU^T = I = U^T U$

Linear Independence and Rank

- A set of vectors $\{x_1, \ldots, x_n\}$ is linearly independent if $\nexists\{\alpha_1, \ldots, \alpha_n\}$: $\sum_{i=1}^{n} \alpha_i x_i = 0$
- Rank: $A \in \mathbb{R}^{m \times n}$, then $rank(A)$ is the maximum number of linearly independent columns (or equivalently, rows)
- Properties:
  - $rank(A) \leq \min\{m, n\}$
  - $rank(A) = rank(A^T)$
  - $rank(AB) \leq \min\{rank(A), rank(B)\}$
  - $rank(A + B) \leq rank(A) + rank(B)$

Example of Linear Dependence

These three vectors are linearly dependent because they all lie in the same plane. The matrix with these three vectors as rows has rank 2.



Rank from row-echelon forms

$$\begin{bmatrix} 1 & 2 & 1 \\ -2 & -3 & 1 \\ 3 & 5 & 0 \end{bmatrix} R_2 \to 2r_1+r_2 \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 3 & 5 & 0 \end{bmatrix} R_3 \to -3r_1+r_3 \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & -1 & -3 \end{bmatrix}$$

$$R_3 \to r_2+r_3 \begin{bmatrix} 1 & 2 & 1 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix} R_1 \to -2r_2+r_1 \begin{bmatrix} 1 & 0 & -5 \\ 0 & 1 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$

Matrix Inversion

- If $A \in \mathbb{R}^{n \times n}$, $rank(A) = n$, then the inverse of A, denoted A^{-1} is the matrix that: $AA^{-1} = A^{-1}A = I$
- Properties:
  - $(A^{-1})^{-1} = A$
  - $(AB)^{-1} = B^{-1}A^{-1}$
  - $(A^{-1})^T = (A^T)^{-1}$
- The inverse of an orthogonal matrix is its transpose

Eigenvalues and Eigenvectors

- Given $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{C}$ is an eigenvalue of A with the corresponding eigenvector $x \in \mathbb{C}^n$ ($x \neq 0$) if:

$$Ax = \lambda x$$

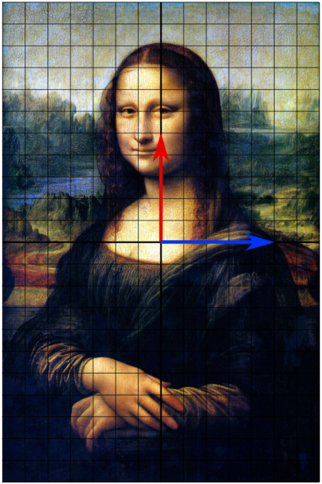
For example, if

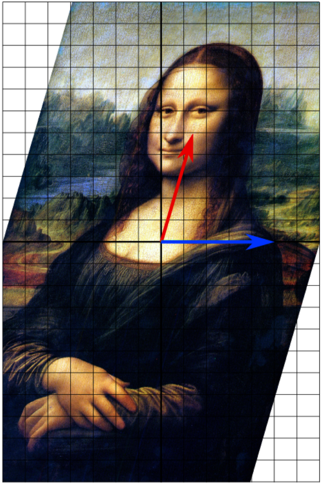
$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

then the vector $\begin{bmatrix} 3 \\ -3 \end{bmatrix}$ is an eigenvector with eigenvalue 1, because

$$A\mathbf{x} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ -3 \end{bmatrix} = \begin{bmatrix} 2 \cdot 3 + 1 \cdot (-3) \\ 1 \cdot 3 + 2 \cdot (-3) \end{bmatrix} = \begin{bmatrix} 3 \\ -3 \end{bmatrix} = 1 \cdot \begin{bmatrix} 3 \\ -3 \end{bmatrix}.$$

Eignevector Example





Eigenvalues and Eigenvectors

- Characteristic Polynomial: If $Ax = \lambda x$ then

$$(A - \lambda I)x = 0$$

so $(A - \lambda I)$ is singular, and we see that

$$\det(A - \lambda I) = 0.$$

Thus the eigenvalues are exactly the n possibly complex roots of the degree n polynomial equation $det(A - \lambda I) = 0$. This polynomial $det(A - \lambda I) = 0$ is known as the characteristic polynomial.

Eigenvalues and Eigenvectors Properties

- Usually eigenvectors are normalized to unit length.
- If A is symmetric, then all the eigenvalues are real and the eigenvectors are orthogonal to each other.
- $tr(A) = \sum_{i=1}^{n} \lambda_i$
- $det(A) = \prod_{i=1}^{n} \lambda_i$
- $rank(A) = |\{1 \leq i \leq n | \lambda_i \neq 0\}|$

Proofs

Induction:

1. Show result on base case, associated with $n = k_0$
2. Assume result true for $n \leq i$. Prove result for $n = i + 1$
3. Conclude result true for all $n \geq k_0$

Proofs

Induction:

1. Show result on base case, associated with $n = k_0$
2. Assume result true for $n \leq i$. Prove result for $n = i + 1$
3. Conclude result true for all $n \geq k_0$

Example:
For all natural number n, $1 + 2 + 3 + ... + n = \frac{n*(n+1)}{2}$
Base case: when $n = 1$, $1 = 1$.
Assume statement holds for $n = k$, then
$1 + 2 + 3 + ... + k = \frac{k*(k+1)}{2}$.
We see $1 + 2 + 3 + ... + (k + 1) = \frac{k*(k+1)}{2} + (k + 1) = \frac{(k+1)(k+2)}{2}$.

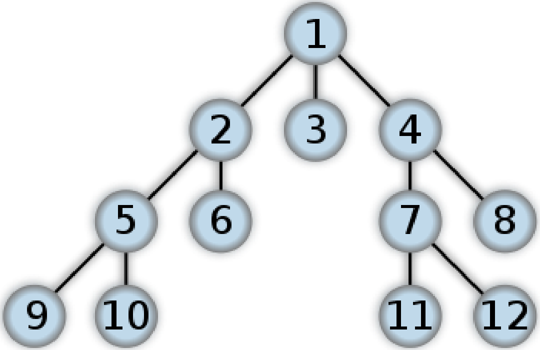
Graph theory

- Definitions: vertex/node, edge/link, loop/cycle, degree, path, neighbor, tree, clique,...
- Random graph (Erdos-Renyi): Each possible edge is present independently with some probability p
- (Strongly) connected component: subset of nodes that can all reach each other
- Diameter: longest minimum distance between two nodes
- Bridge: edge connecting two otherwise disjoint connected components

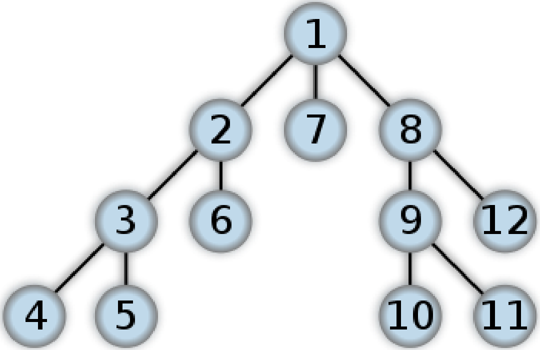
Basic algorithms

- BFS: explore by “layers”
- DFS: go as far as possible, then backtrack
- Greedy: maximize goal at each step
- Binary search: on ordered set, discard half of the elements at each step

Breadth First Search

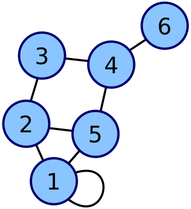


Depth First Search



Adjacency Matrix

The adjacency matrix M of a graph is the matrix such that $M_{i,j} = 1$ if i is connected to j, and $M_{i,j} = 0$ otherwise.



$$\begin{bmatrix} 1 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 1 & 1 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}$$

For example, this is useful when studying random walks. Renormalize the rows of M so that every row has sum 1. Then if we start at vertex i, after k random walk steps, the distribution of our location is $M^k e_i$, where e_i has a 1 in the ith coordinate and 0 elsewhere.