Linear Algebra
Prerequisites - continued

Jana Kosecka
http://cs.gmu.edu/~kosecka/cs682.html
kosecka@cs.gmu.edu
Matrices

- **n x m matrix**
  - meaning:
  - m points from n-dimensional space
  - transformation

- **Covariance matrix** - symmetric
  - Square matrix associated with
  - The data points (after mean has been subtracted) in 2D

\[ C = A A^T \]

- **Special case**
  - matrix is square

\[ A \in \mathbb{R}^{2\times 2} \]

\[ y = Ax \]
Geometric interpretation

Lines in 2D space - row solution
Equations are considered isolation

\[ 2x - y = 1 \]
\[ x + y = 5 \]

Linear combination of vectors in 2D
Column solution

\[
\begin{bmatrix}
2 \\
1
\end{bmatrix} x + \begin{bmatrix}
-1 \\
1
\end{bmatrix} y = \begin{bmatrix}
1 \\
5
\end{bmatrix}
\]

We already know how to multiply the vector by scalar
Linear equations

In 3D

\[
\begin{bmatrix}
2 & 1 & 1 \\
4 & -6 & 0 \\
-2 & 7 & 2
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
w
\end{bmatrix}
= \begin{bmatrix}
5 \\
-2 \\
9
\end{bmatrix}
\]

When is RHS a linear combination of LHS

\[
\begin{bmatrix}
2 \\
4 \\
-2
\end{bmatrix}u
+ \begin{bmatrix}
1 \\
-6 \\
7
\end{bmatrix}v
+ \begin{bmatrix}
1 \\
0 \\
2
\end{bmatrix}w
= \begin{bmatrix}
5 \\
-2 \\
9
\end{bmatrix}
\]

Solving linear n equations with n unknowns

If matrix is invertible - compute the inverse

Columns are linearly independent

\[Ax = y\]
\[det(A) \neq 0\]
\[A^{-1}Ax = A^{-1}y\]
\[x = A^{-1}y\]
Linear equations

Not all matrices are invertible

- inverse of a 2x2 matrix (determinant non-zero)
- inverse of a diagonal matrix

Computing inverse - solve for the columns
Independently or using Gauss-Jordan method

\[
\begin{bmatrix}
2 & 1 & 1 \\
4 & -6 & 0 \\
-2 & 7 & 2
\end{bmatrix}
\begin{bmatrix}
x_1 \\
x_2 \\
x_3
\end{bmatrix}
= 
\begin{bmatrix}
1 & 0 & 0 \\
0 & 1 & 0 \\
0 & 0 & 1
\end{bmatrix}
\]
Vector spaces (informally)

• Vector space in n-dimensional space \( \mathbb{R}^n \)
• n-dimensional columns with real entries
• Operations of addition, multiplication and scalar multiplication
• Additions of the vectors and multiplication of a vector by a scalar always produces vectors which lie in the space

• Matrices also make up vector space - e.g. consider all 3x3 matrices as elements of \( \mathbb{R}^9 \) space
Vector subspace

A subspace of a vector space is a non-empty set of vectors closed under vector addition and scalar multiplication.

Example: overconstrained system - more equations than unknowns

\[
\begin{bmatrix}
  u_1 & v_1 \\
  u_2 & v_2 \\
  u_3 & v_3 \\
\end{bmatrix}
\begin{bmatrix}
  x_1 \\
  x_2 \\
\end{bmatrix}
= 
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3 \\
\end{bmatrix}
\]

The solution exists if \( b \) is in the subspace spanned by vectors \( u \) and \( v \)

\[
\begin{bmatrix}
  u_1 \\
  u_2 \\
  u_3 \\
\end{bmatrix} x_1 + 
\begin{bmatrix}
  v_1 \\
  v_2 \\
  v_3 \\
\end{bmatrix} x_2 = 
\begin{bmatrix}
  b_1 \\
  b_2 \\
  b_3 \\
\end{bmatrix}
\]
Linear Systems - Nullspace

1. When matrix is square and invertible
2. When the matrix is square and noninvertible
3. When the matrix is non-square with more constraints than unknowns

\[ Ax = b \]

Solution exists when \( b \) is in column space of \( A \)

Special case

All the vectors which satisfy \( Ax = 0 \) lie in the NULLSPACE of matrix \( A \)
Basis

n x n matrix $A$ is invertible if it is of a full rank

Rank of the matrix - number of linearly independent rows (see definition next page)

If the rows of columns of the matrix $A$ are linearly independent - the nullspace of contains only $0$ vector

Set of linearly independent vectors forms a basis of the vector space

Given a basis, the representation of every vector is unique

Basis is not unique (examples)
Linear independence

**Definition A.1 (A linear space).** A set (of vectors) $V$ is considered as a linear space over the field $\mathbb{R}$, if its elements, so-called vectors, are closed under two basic operations: scalar multiplication and vector summation. That is, given any two vectors $v_1, v_2 \in V$ and any two scalars $\alpha, \beta \in \mathbb{R}$, the linear combination $v = \alpha v_1 + \beta v_2$ is also a vector in $V$.

**Definition A.4 (Linearly independence).** A set of vectors $S = \{v_i\}_{i=1}^m$ is said to be linearly independent if the equation

$$\alpha_1 v_1 + \alpha_2 v_2 + \cdots + \alpha_m v_m = 0$$

implies

$$\alpha_1 = \alpha_2 = \cdots = \alpha_m = 0.$$

**Definition A.5 (Basis).** A set of vectors $B = \{b_i\}_{i=1}^n$ of a linear space $V$ is said to be a basis if $B$ is a linearly independent set and $B$ spans the entire space $V$ (i.e. $V = \text{span}(B)$).
Change of basis

Fact A.6 (Properties of basis). Suppose $B$ and $B'$ are two bases for a linear space $V$. Then

2. Let $B = \{b_i\}_{i=1}^{n}$ and $B' = \{b'_i\}_{i=1}^{n}$, then each base vector of $B$ can be expressed as linear combination of those in $B'$, i.e.

$$b_j = a_{1j}b'_1 + a_{2j}b'_2 + \cdots + a_{nj}b'_n = \sum_{i=1}^{n} a_{ij}b'_i. \quad (A.2)$$

for some $a_{ij} \in \mathbb{R}, i, j = 1, 2, \ldots, n$.

3. For any vector $v \in V$, it can be written as a linear combination of vectors in either of the bases

$$v = x_1b_1 + x_2b_2 + \cdots + x_nb_n = x'_1b'_1 + x'_2b'_2 + \cdots + x'_nb'_n \quad (A.3)$$

where the coefficients $\{x_i \in \mathbb{R}\}_{i=1}^{n}$ and $\{x'_i \in \mathbb{R}\}_{i=1}^{n}$ are uniquely determined and are called the coordinates of $v$ with respect to each basis.
Change of basis (contd.)

\[
[b_1, b_2, \ldots, b_n] = [b'_1, b'_2, \ldots, b'_n] \begin{bmatrix}
    a_{11} & a_{12} & \cdots & a_{1n} \\
    a_{21} & a_{22} & \cdots & a_{2n} \\
    \vdots & \vdots & \ddots & \vdots \\
    a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}.
\]

\[
v = [b_1, b_2, \ldots, b_n]\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix} = [b'_1, b'_2, \ldots, b'_n]\begin{bmatrix}
a_{11} & a_{12} & \cdots & a_{1n} \\
a_{21} & a_{22} & \cdots & a_{2n} \\
\vdots & \vdots & \ddots & \vdots \\
a_{n1} & a_{n2} & \cdots & a_{nn}
\end{bmatrix}\begin{bmatrix}
x_1 \\
x_2 \\
\vdots \\
x_n
\end{bmatrix}
\]

\[
B^t = BA^{-1}, \quad x^t = Ax.
\]
Linear Equations

In general

Four basic subspaces

- Column space of $A$ - dimension of $C(A)$
  number of linearly independent columns
  $r = \text{rank}(A)$

- Row space of $A$ - dimension of $R(A)$
  number of linearly independent rows
  $r = \text{rank}(A^T)$

- Null space of $A$ - dimension of $N(A)$ $n - r$

- Left null space of $A$ - dimension of $N(A^T)$ $m - r$

- Maximal rank - min($n,m$) - smaller of the two dimensions

Vector space spanned by columns of $A$

\[
\begin{bmatrix}
2 & 1 & 1 \\
4 & -6 & 0 \\
-2 & 7 & 2 \\
\end{bmatrix}
\begin{bmatrix}
u \\
v \\
w
\end{bmatrix}
= 
\begin{bmatrix}
5 \\
-2 \\
9
\end{bmatrix}
\]
Linear Equations

Vector space spanned by columns of $A$

$\begin{bmatrix} 2 \\ 4 \\ -2 \end{bmatrix} u + \begin{bmatrix} 1 \\ -6 \\ 7 \end{bmatrix} v + \begin{bmatrix} 1 \\ 0 \\ 2 \end{bmatrix} w = \begin{bmatrix} 5 \\ -2 \\ 9 \end{bmatrix}$

In general

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

Four basic possibilities, suppose that the matrix $A$ has full rank

Then:

- if $n < m$ number of equations is less then number of unknowns, the set of solutions is $(m-n)$ dimensional vector subspace of $\mathbb{R}^m$
- if $n = m$ there is a unique solution
- if $n > m$ number of equations is more then number of unknowns, there is no solution
Structure induced by a linear map

$x_r \rightarrow Ax_r = Ax$

$x \rightarrow Ax$

$Ax_n = 0$
Linear Equations - Square Matrices

1. \( A \) is square and invertible
2. \( A \) is square and non-invertible

1. System \( Ax = b \) has at most one solution - columns
   are linearly independent \( \text{rank} = n \)
   - then the matrix is invertible \( x = A^{-1}y \)
2. Columns are linearly dependent \( \text{rank} < n \)
   - then the matrix is not invertible
Linear Equations – non-square matrices

Long-tin matrix
over-constrained system

\[
\begin{bmatrix}
2 \\
3 \\
4
\end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}
\]

\(a x = b\)

The solution exist when \(b\) is aligned with \([2,3,4]^T\)

If not we have to seek some approximation – least squares

Approximation – minimize squared error

\[e^2 = (2x - b_1)^2 + (3x - b_2)^2 + (4x - b_3)^2\]

Least squares solution - find such value of \(x\) that the error

Is minimized (take a derivative, set it to zero and solve for \(x\))

Short for such solution

\[e^2 = \|ax - b\|^2\]

\[\begin{align*}
a x &= b \\
a^T ax &= a^T b \\
\bar{x} &= \frac{a^T b}{a^T a}
\end{align*}\]
Linear equations – non-squared matrices

Similarly when $A$ is a matrix

\[
\begin{bmatrix}
1 & 2 \\
1 & 3 \\
0 & 0
\end{bmatrix}
\begin{bmatrix}
x \\
5
\end{bmatrix}
= 
\begin{bmatrix}
4 \\
6
\end{bmatrix}
\]

\[e^2 = \|Ax - b\|^2\]

\[Ax = b\]

\[A^T Ax = A^T b\]

\[\bar{x} = (A^T A)^{-1} A^T b\]

\[A^\dagger = (A^T A)^{-1} A^T\]

is so called pseudoinverse of matrix $A$
Homogeneous Systems of equations

\[ Ax = 0 \]

When matrix is square and non-singular, there a Unique trivial solution \( x = 0 \)

If \( m \geq n \) there is a non-trivial solution when rank of \( A \) is \( \text{rank}(A) < n \)
We need to impose some constraint to avoid trivial Solution, for example
\[
\| x \| = 1
\]

Find such \( x \) that \( \| Ax \|^2 \) is minimized
\[
\| Ax \|^2 = x A^T A x
\]

Solution: eigenvector associated with the smallest eigenvalue
Eigenvalues and Eigenvectors

• Motivated by solution to differential equations
• For square matrices $A \in \mathbb{R}^{n \times n}$ $\dot{u} = Au$

For scalar ODE's

$\dot{u} = au$
$u(t) = e^{at}u(0)$

We look for the solutions of the following type exponentials

$v(t) = e^{\lambda t}y$
$w(t) = e^{\lambda t}z$

Substitute back to the equation

$\lambda e^{\lambda t}y = 4e^{\lambda t}y - 5e^{\lambda t}z$
$\lambda e^{\lambda t}z = 2e^{\lambda t}y - 3e^{\lambda t}z$

$x = \begin{bmatrix} y \\ z \end{bmatrix}$
$\lambda x = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} x$
Eigenvalues and Eigenvectors

\[ \lambda x = \begin{bmatrix} 4 & -5 \\ 2 & -3 \end{bmatrix} x \]

\[ Ax = \lambda x \]

Solve the equation:

\[ (A - \lambda I)x = 0 \quad (1) \]

\( x \) - is in the null space of \( (A - \lambda I) \)

\( \lambda \) is chosen such that \( (A - \lambda I) \) has a null space

Computation of eigenvalues and eigenvectors (for dim 2,3)
1. Compute determinant
2. Find roots (eigenvalues) of the polynomial such that determinant = 0
3. For each eigenvalue solve the equation (1)

For larger matrices - alternative ways of computation
Eigenvalues and Eigenvectors

For the previous example

\[ \lambda_1 = -1, \quad x_1 = [1, 1]^T \quad \lambda_2 = -2, \quad x_2 = [5, 2]^T \]

We will get special solutions to ODE

\[ \dot{u} = Au \]

\[ Au = e^{\lambda_1 t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} \quad u = e^{\lambda_2 t} \begin{bmatrix} 5 \\ 2 \end{bmatrix} \]

Their linear combination is also a solution (due to the linearity of \( \dot{u} = Au \))

\[ u = c_1 e^{\lambda_1 t} \begin{bmatrix} 1 \\ 1 \end{bmatrix} + c_2 e^{\lambda_1 t} \begin{bmatrix} 5 \\ 2 \end{bmatrix} \]

In the context of diff. equations – special meaning
Any solution can be expressed as linear combination
Individual solutions correspond to modes
Eigenvalues and Eigenvectors

\[ Ax = \lambda x \]

Only special vectors are eigenvectors
- such vectors whose direction will not be changed by the transformation \( A \) (only scale)
- they correspond to normal modes of the system act independently

Examples

\[
A = \begin{bmatrix}
2 & 0 \\
0 & 3
\end{bmatrix}
\]

\[
\begin{align*}
eigenvalues & \quad eigenvectors \\
2, 3 & \quad \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix}
\end{align*}
\]

Whatever \( A \) does to an arbitrary vector is fully determined by its eigenvalues and eigenvectors

\[ Ax = 2\lambda_1 v_1 + 5\lambda_2 v_2 \]
Eigenvalues and Eigenvectors - Diagonalization

• Given a square matrix $A$ and its eigenvalues and eigenvectors - matrix can be diagonalized

$$A = S \Lambda S^{-1} \quad A = S \Lambda S^{-1}$$

Matrix of eigenvectors $\rightarrow$ Diagonal matrix of eigenvalues

$$A S = \Lambda S$$

$$A\begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} = \begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \ldots & \lambda_n x_n \end{bmatrix} \quad Ax = \lambda x$$

$$\begin{bmatrix} \lambda_1 x_1 & \lambda_2 x_2 & \ldots & \lambda_n x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & \lambda_2 & \ldots & \lambda_n \end{bmatrix}$$

$$A = S \Lambda S^{-1}$$

• If some of the eigenvalues are the same, eigenvectors are not independent
Diagonalization

- If there are no zero eigenvalues - matrix is invertible
- If there are no repeated eigenvalues - matrix is diagonalizable
- If all the eigenvalues are different then eigenvectors are linearly independent

For Symmetric Matrices

If $A$ is symmetric \[ A = Q \Lambda Q^T \]

orthonormal matrix of eigenvectors

Diagonal matrix of eigenvalues

i.e. for a covariance matrix or some matrix $B = A^T A$
Symmetric matrices (contd.)

\[ A^T \tilde{A} = V \text{diag}\{\sigma_1^2, \sigma_2^2, \ldots, \sigma_n^2\} V^T \]

\[ \| A \|_f = \sqrt{\sum_{i,j} a_{i,j}^2}. \]

\[ \| A \|_f = \sqrt{\text{trace}(A^T A)} = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}. \]
Example - line fitting

Equation of a line
\[ ax + by = d \]

Line normal
\[ \mathbf{n} = [a, b] \]

Distance to the origin
\[ d \]

Error function
\[ e(a, b, d) = \sum_{i=1}^{n} (ax_i + by_i - d)^2 \]

Differentiate with respect to \( a, b, d \)
set the first derivative to 0 and solve for the parameters
Least squares line fitting

• Data: \((x_1, y_1), \ldots, (x_n, y_n)\)
• Line equation: \(y_i = mx_i + b\)
• Find \((m, b)\) to minimize

\[
E = \sum_{i=1}^{n} (y_i - mx_i - b)^2
\]

\[
\begin{bmatrix}
  y_1 \\
  \vdots \\
  y_n
\end{bmatrix}
= \begin{bmatrix}
  x_1 & 1 \\
  \vdots & \vdots \\
  x_n & 1
\end{bmatrix}
\begin{bmatrix}
  m \\
  b
\end{bmatrix}
\]

\[
E = \left\| \vec{b} - A\vec{x} \right\|^2 = (\vec{b} - A\vec{x})^T (\vec{b} - A\vec{x}) = \vec{b}^T \vec{b} - 2(A\vec{x})^T \vec{b} + (A\vec{x})^T (A\vec{x})
\]

\[
\frac{dE}{d\vec{x}} = 2A^T A\vec{x} - 2A^T \vec{b} = 0
\]

\[
A^T A\vec{x} = A^T \vec{b}
\]

Normal equations: least squares solution to

\[
A\vec{x} = \vec{b}
\]
Problem with “vertical” least squares

- Not rotation-invariant
- Fails completely for vertical lines
Total least squares

• Distance between point
• \((x_i, y_i)\) and line \(ax + by = d\)
\((a^2 + b^2 = 1)\): \(|ax_i + by_i - d|\)

Unit normal: \(N = (a, b)\)
Total least squares

- Distance between point \((x_i, y_i)\) and line \(ax + by = d\) \((a^2 + b^2 = 1)\):
  \[|ax_i + by_i - d|\]

- Find \((a, b, d)\) to minimize the sum of squared perpendicular distances

\[
E = \sum_{i=1}^{n} (ax_i + by_i - d)^2
\]
Total least squares

- Distance between point \((x_i, y_i)\) and line \(ax+by=d\) \((a^2+b^2=1)\): 
  \[|ax_i + by_i - d|\]

- Find \((a, b, d)\) to minimize the sum of squared perpendicular distances

\[E = \sum_{i=1}^{n} (ax_i + by_i - d)^2\]

\[
\frac{\partial E}{\partial d} = \sum_{i=1}^{n} -2(ax_i + by_i - d) = 0
\]

\[E = \sum_{i=1}^{n} (a(x_i - \bar{x}) + b(y_i - \bar{y}))^2 = \left\| \begin{bmatrix}
  x_1 - \bar{x} & y_1 - \bar{y} \\
  \vdots & \vdots \\
  x_n - \bar{x} & y_n - \bar{y}
\end{bmatrix} \begin{bmatrix}
  a \\
  b
\end{bmatrix} \right\|^2 = (A\tilde{u})^T (A\tilde{u})\]

\[
\tilde{u} = \begin{bmatrix}
  a \\
  b
\end{bmatrix} \quad \frac{dE}{d\tilde{u}} = 2(A^T A)\tilde{u} = 0
\]
Total least squares

Solution to \((AT^A)u = 0\), subject to \(||u||^2 = 1\): eigenvector of \(AT^A\) associated with the smallest eigenvalue (least squares solution to homogeneous linear system \(A\vec{u} = 0\))

In case of 2D line fitting

\[
A = \begin{bmatrix}
x_1 - \bar{x} & y_1 - \bar{y} \\
\vdots & \vdots \\
x_n - \bar{x} & y_n - \bar{y}
\end{bmatrix}, \quad \quad AT^A = \begin{bmatrix}
\sum_{i=1}^{n}(x_i - \bar{x})^2 & \sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) \\
\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^{n}(y_i - \bar{y})^2
\end{bmatrix}
\]

second moment matrix - geometric interpretation of eigenvalues and eigenvectors
\[
A^T A = \begin{bmatrix}
\sum_{i=1}^{n} (x_i - \bar{x})^2 & \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \\
\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) & \sum_{i=1}^{n} (y_i - \bar{y})^2
\end{bmatrix}
\]

second moment matrix