**MASSACHUSETTS INSTITUTE OF TECHNOLOGY**
**Department of Civil and Environmental Engineering**

**1.731 Water Resource Systems**

---

**Lecture 2, Linear Algebra Review, Sept. 12, 2006**

---

The notation and some of the basic concepts of linear algebra are needed in optimization theory, which is concerned with large systems of equations in many variables.

You should learn or review the following topics. See any introductory linear algebra text for details.

**Indicial notation:** **Vector**: $x = [x_1, x_2, \ldots, x_n] \rightarrow x_i$

**Matrix**: $A = \begin{bmatrix} A_{11} & \cdots & A_{1n} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mn} \end{bmatrix} \rightarrow A_{ij}$

**Matrix transpose and symmetry:**
MATLAB operator ′

$$A^T \rightarrow A_{ij}^T = A_{ji}$$

$A$ is **symmetric** if $A^T = A$ (no change if rows and columns are interchanged)

**Vector & matrix operations products:**
MATLAB operators + and *

$z = x + y \rightarrow z_i = x_j + y_i$ **Vector sum**

$C = A + B \rightarrow C_{ij} = A_{ij} + B_{ij}$ **Matrix sum**

$z = ax \rightarrow z_i = ax_j \quad B = aA \rightarrow B_{ij} = aA_{ij}$ **Scalar multiplication**

$z = x^T y \rightarrow z = x_j y_j$ **Scalar product,** $x$ and $y$ are **orthogonal** if $x^T y = 0$

$y = Ax \rightarrow y_i = A_{ij} x_j$ **Matrix-vector product**, implied sum over repeated indices ($j$)

$C = AB \rightarrow C_{ik} = A_{ij} B_{jk}$ **Matrix product**, implied sum over repeated indices ($j$)

$q = x^T Ax \rightarrow q = x_i A_{ij} x_j$ **quadratic form**, implied sum over repeated indices ($i,j$)

**Systems of linear equations:**

$$Ax = b \rightarrow A_{ij}x_j = b_i \quad m \text{ equations, } n \text{ unknowns } x_j \ (i = 1,\ldots, m \ \ j = 1,\ldots,n)$$

**Row echelon form of a matrix**

This is convenient for analyzing and solving systems of linear equations.

A is in **row echelon form** if:

- All rows with non-zero entries are above rows containing only zeros
- Leading (non-zero) entry of each row is to right of leading entry in row above it
- All entries below a leading coefficient are zero

*A* is **row echelon**:

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 9/2 \end{bmatrix}$$

*A* is **not row echelon**

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{bmatrix}$$

**Reduced row echelon form of a matrix**

MATLAB function rref(A)

Add requirements that:

- All leading entries = 1
- All entries above and below leading entries = 0

*A* is **reduced row echelon**:

$$A = \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 9/2 \end{bmatrix}$$

**Using elementary row operations to derive row and reduced row echelon forms:**

- **Elementary row operations** $\rightarrow$ replace a given row with weighted sum of any two rows.
- Carry out elementary row operations starting from top row down to meet requirements for **row echelon form**.

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 3 & 4 & 6 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 9/2 \end{bmatrix}$$

$$(3/2)r_1 + (-1/2)(r_2) \rightarrow r_2$$

- Carry out additional elementary row operations starting from top row down to meet requirements for **reduced row echelon form**.

$$A = \begin{bmatrix} 1 & 2 & 5 \\ 0 & 1 & 9/2 \end{bmatrix} \quad \rightarrow \quad \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 9/2 \end{bmatrix}$$

$$r_1 + (-2)(r_2) \rightarrow r_1$$

**Gaussian elimination:**

MATLAB operator x=A\b

This is a procedure for solving systems of linear equations by applying a series of elementary row operations:

- Augment *A* by appending *b* as last column to obtain $[A \mid b]$
- Put $[A \mid b]$ in **reduced row echelon** form

- Solution $x$ is last column of row echelon matrix

$$Ax = b \quad A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad b = \begin{bmatrix} 5 \\ 6 \end{bmatrix}$$

$$[A\,|\,b] = \left[\begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array}\right] \rightarrow \begin{bmatrix} 1 & 0 & -4 \\ 0 & 1 & 9/2 \end{bmatrix} \qquad x = \begin{bmatrix} -4 \\ 9/2 \end{bmatrix}$$

**Matrix rank**

MATLAB function rank(A)

**Rank** of a matrix is number of non-zero rows in row echelon form:

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 2 & 3 \\ 0 & 0 & 0 \end{bmatrix} \qquad \text{Rank = number of non-zero rows = 1}$$

**Consistency and uniqueness**

- System of linear equations $A\underline{x} = \underline{b}$ is **consistent** if $Rank(A) = Rank([A\,|\,b])$
- The $m$ by $n$ homogeneous system $Ax$=0 always has the **trivial solution** $x = 0$.
- An $m$ by $n$ consistent non-homogeneous system $Ax$=$b$ has a **unique solution** if $Rank(A)$ = $n$ = number of unknowns.
- An $m$ by $n$ consistent non-homogeneous system $Ax$=$b$ has a **non-trivial non-unique solution** if $Rank(A) = r < n$ = number of unknowns. The number of **free parameters** in the solution is $n - r$.

**Determinant of a square matrix:**

MATLAB function det(A)

Determinant $|A|$ of $A$ is a scalar matrix property useful for solving eigen problems.

Determinant can be evaluated from row echelon form (which is upper triangular for a square matrix). Apply the following rules:

- If an elementary row operation that transforms $A$ to $B$ has the form $c_i r_i + c_j r_j \rightarrow r_j$, then $|A| = |B| / c_j$.
- $|A|$= product of the leading (diagonal) terms of the final row echelon form.

In this example row echelon is produced by an elementary row operation that replaces row 2 using $c_2$ =1:

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \rightarrow \begin{vmatrix} 1 & 2 \\ 0 & -2 \end{vmatrix} = (-2)/c_2 = (-2)/1 = -2 \qquad \begin{array}{l} (-3)r_1 + r_2 \rightarrow r_2 \\ c_2 = 1 \end{array}$$

In this example row echelon is produced by an elementary row operation that replaces row 2 using $c_2$ =-1/2:

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} \rightarrow \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} = (1)/c_2 = (1)/(-1/2) = -2 \qquad \begin{matrix} (3/2)r_1 + (-1/2)(r_2) \rightarrow r_2 \\ c_2 = -1/2 \end{matrix}$$

**Inverse of a square matrix:**
MATLAB function inv(A)

$$A^{-1}A = AA^{-1} = I \quad \rightarrow \quad \lfloor A^{-1} \rfloor_{ij} A_{jk} = \delta_{ik}$$

Find $A^{-1}$ by solving $A[A^{-1}] = I$ using Gaussian elimination (with $A$ given and each column of $A^{-1}$ considered to be an unknown vector).

If $Rank(A) = r < n$ then $|A|=0$ and matrix is **singular** and has no inverse

**Linear Dependence/Independence**

**Linear combination** of a set of $m$ $n$-dimensional vectors $x_{1j}, \dots, x_{mj}$ is:

$$\sum_{j=1}^{m} a_j x_{ij} = a_j x_{ij}$$ ; $x_{ij}$ is an $n$ by $m$ matrix whose columns are the $x_{i1}, \dots, x_{jim}$ vectors

The vectors $x_{j1}, \dots, x_{jm}$ are **linearly independent** if $a_j x_{ij}$ has a unique (trivial) solution $a_j = 0$.
If the $a_j$*'s* can be non-zero the vectors are **linearly dependent**.

**Linear Vector Spaces, Subspaces, and Projections**

The set of all possible $n$-dimensional vectors $[x_1, x_2 \dots x_n]$ forms a **linear vector space** $V_n$ since it is closed under vector addition and scalar multiplication (i.e. any vector $a_1 x_{i1} + a_2 x_{i2}$ is in $V_n$ if the vectors $x_{i1}$ and $x_{i2}$ are in $V_n$).

Consider the $n$ by $m$ matrix $A_{ij}$ with columns consisting of the $m$ vectors $A_{i1}, \dots, A_{im}$ from $V_n$. The set of vectors that are linear combinations of these $m$ column vectors form a linear vector space $V_m$ which is a **subspace** (subset) of $V_n$. The subspace $V_m$ **is spanned** by the $A_{i1}, \dots, A_{im}$. If the $m$ spanning vectors are linearly independent they form a **basis** for $V_m$. The **dimension** of $V_m$ is the rank of $A_{ij}$, which will be equal to $m$ if the spanning vectors are linearly independent and form a basis. If $m = n = Rank\,(A)$ then $V_m = V_n$.

If the $m$ columns of $A_{ij}$ are linearly independent the $n$-$m$ solutions of the system $A_{ij}\, x_i = 0$ form a basis for an $n$-$m$ dimensional subspace $V_{n-m}$ of $V_n$. $V_{n-m}$ is the **null space** of $V_m$ and vice versa. Every vector in $V_{n-m}$ is orthogonal to every vector in $V_m$.

Each basis vector of $V_m$ may be viewed as the **normal vector** to a $n$-1 dimensional **hyperplane**. The intersection of all $m$ such hyperplanes is an $n$-$m$ hyperplane that contains all vectors in the null space $V_{n-m}$.

The **projection** $P_{ij}x_j$ of any vector $x_i$ in $V_n$ onto the subspace $V_m$ is a vector that 1) lies in $V_m$ and 2) obeys the property $x_i = P_{ij}x_j + x_{\perp i}$ , where $x_{\perp i}$ lies in the null space $V_{n\text{-}m}$ .(i.e. $x_i$ - $P_{ij}x_j$ is orthogonal to all the vectors in $V_m$). It follows from these properties that the $n$ by $n$ **projection matrix** $P_{ij}$ is:

$$P = A[A^T A]^{-1} A^T$$

The matrix $m$ by $m$ $[A^T A]$ is invertible since $A$ has rank $m$.

**Eigen problems**

Eigen problem seeks a set of scalar **eigenvalues** $\lambda^k$ and **eigenvectors** $u_k$, for $k = 1,\ldots, n$ associated with the $n$ by $n$ matrix $A$.

The eigenvalues and eigenvectors satisfy:

$$A_{ij}u_j^k = \lambda^k u_i^k \qquad k = 1,\ldots n \qquad \lambda^k \text{ and } u_i^k \text{ are unknown}$$

The $n$ eigenvalues are found by solving $n$th order polynomial in $\lambda$ for $n$ roots :

$$|A - \lambda I| = 0 \quad \rightarrow \quad \lambda^1, \lambda^2, \ldots, \lambda^n$$

The corresponding $n$ eigenvectors are found by substituting each $\lambda^k$ into $A_{ij}u_j^k = \lambda^k u_i^k$ and solving for the corresponding $u_i^k$ .

Example: $A = \begin{bmatrix} 1 & 0 \\ 3 & 4 \end{bmatrix}$

Eigenvalues: $\begin{vmatrix} 1-\lambda & 0 \\ 3 & 4-\lambda \end{vmatrix} = (1-\lambda)(4-\lambda) \rightarrow \lambda^1 = 1, \lambda^2 = 4$

Eigenvector 1: $\begin{bmatrix} 1-\lambda^1 & 0 \\ 3 & 4-\lambda^1 \end{bmatrix}\begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 3 & 3 \end{bmatrix}\begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} \rightarrow \begin{bmatrix} u_1^1 \\ u_2^1 \end{bmatrix} = \begin{bmatrix} a \\ -a \end{bmatrix}$ for any $a$

Eigenvector 2: $\begin{bmatrix} 1-\lambda^2 & 0 \\ 3 & 4-\lambda^2 \end{bmatrix}\begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} = \begin{bmatrix} -3 & 0 \\ 3 & 0 \end{bmatrix}\begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} \rightarrow \begin{bmatrix} u_1^2 \\ u_2^2 \end{bmatrix} = \begin{bmatrix} 0 \\ a \end{bmatrix}$ for any $a$

**Quadratic Forms/Definiteness:**

Quadratic form $q(x) = x^T Ax \quad \rightarrow q = x_i A_{ij} x_j$ (or the matrix A that it depends upon) can be classified as follows:

- $q(x)$, $A$ **positive definite** if $q(x) > 0$ for all - if $A$ symmetric **all eigenvalues of $A > 0$**
- $q(x)$, $A$ **positive semidefinite** if $q(x) \geq 0$ for all $x$ - if $A$ symmetric **all eigenvalues of $A \geq 0$**
- $q(x)$, $A$ **negative definite** if $q(x) < 0$ for all $x$ - if $A$ symmetric **all eigenvalues of $A < 0$**
- $q(x)$, $A$ **negative semi-definite** if $q(x) \leq 0$ for all $x$ - if $A$ symmetric **all eigenvalues of $A \leq 0$**
- otherwise $q(x)$ and $A$ are **indefinite**