CSci 6974 and ECSE 6966 Math. Tech. for Vision, Graphics and Robotics
Lecture 2, January 23, 2006
Matrices and Matrix Operations

Late Policy

Homework is due in class on the scheduled due date. If the homework is turned in after class it is late. Students may use up to 4 late days on assignments during the entire semester. From the end of class on the due date until noon the following day counts as one day late. Noon on the following day until noon on the second day is two days, etc.

Notation

- Matrices are rectangular arrays of numbers, with each number subscripted by two indices:

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ . & . & & . \\ . & . & & . \\ . & . & & . \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \tag{1}$$

- A short-hand notation for this is

$$\mathbf{A} = (a_{ij}). \tag{2}$$

- The dimensions of this matrix are written $m \times n$ for m rows and n columns.
- Matrices will be denoted with bold capital letters, generally taken from the beginning and middle of the alphabet, e.g.

$$\mathbf{A}, \mathbf{K}, \mathbf{M}. \tag{3}$$

- The $n \times n$ identity matrix, denoted by $\mathbf{I}_n$ or $\mathbf{I}_{n\times n}$ when the the value of n must be made clear, is a square matrix such that:

$$\mathbf{I}_n = (\delta_{ij}) \tag{4}$$

  where

$$\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases} \tag{5}$$

  is the Kronecker delta.

- Vectors will generally be treated as column vectors ($m \times 1$ matrices) instead of row vectors ($1 \times n$ matrices).

Matrix Operations — Transpose and Symmetry

- The *transpose* of a matrix is obtained by interchanging its rows and columns. It is denoted with the superscript T, as in

$$\mathbf{A}^\top. \tag{6}$$

- Note that $(\mathbf{A}^\top)^\top = \mathbf{A}$.
- When $\mathbf{A}^\top = \mathbf{A}$, then $\mathbf{A}$ is said to be *symmetric*.

Matrix Operations — Addition and Multiplication

- The *sum* of two $m \times n$ matrices $\mathbf{A}$ and $\mathbf{B}$ is simply

$$(a_{ij} + b_{ij}). \tag{7}$$

- The product of a constant c and a matrix $\mathbf{A}$ is

$$c\mathbf{A} = \mathbf{A}c = (ca_{ij}). \tag{8}$$

- The *product* of an $m \times n$ matrix $\mathbf{A}$ and an $n \times p$ matrix $\mathbf{B}$ is an $m \times p$ matrix:

$$\mathbf{AB} = \left(\sum_{k=1}^{n} a_{ik} b_{kj} \right) \tag{9}$$

- Matrix multiplication is associative but not commutative. It distributes over matrix addition (for appropriately sized matrices):

$$\mathbf{A}(\mathbf{B} + \mathbf{C}) = \mathbf{AB} + \mathbf{AC} \tag{10}$$

- Note in general that

$$(\mathbf{AB})^\top = \mathbf{B}^\top \mathbf{A}^\top. \tag{11}$$

- **Hint:** Whenever doing matrix multiplication (as part of a larger set of matrix manipulations), use the fact that the number of columns in the first matrix must equal the number of rows of the second matrix as a sanity check.

More on Matrix Multiplication

- In matrix notation, the dot product of two vectors may be written as

$$\mathbf{u} \cdot \mathbf{v} = \mathbf{u}^\top \mathbf{v} = \mathbf{v}^T \mathbf{u}. \tag{12}$$

- The *outer product* of an $m \times 1$ vector $\mathbf{u}$ and a $n \times 1$ vector $\mathbf{v}$ is a $m \times n$ matrix:

$$\mathbf{u}\mathbf{v}^\top. \tag{13}$$

- The squared-magnitude of a vector $\mathbf{x}$ is

$$\|\mathbf{x}\|^2 = \mathbf{x}^\top \mathbf{x} \tag{14}$$

- The multiplication of a $m \times n$ matrix $\mathbf{A}$ times a $n \times 1$ vector $\mathbf{b}$ may be expressed concisely when $\mathbf{A}$ is written in terms of a stack of row vectors:

$$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \cdots \\ \mathbf{a}_m^\top \end{pmatrix} \tag{15}$$

  where each $\mathbf{a}_i^\top$ contains n entries. Then

$$\mathbf{A}\mathbf{b} = \begin{pmatrix} \mathbf{a}_1^\top \mathbf{b} \\ \mathbf{a}_2^\top \mathbf{b} \\ \cdots \\ \mathbf{a}_m^\top \mathbf{b} \end{pmatrix} \tag{16}$$

- The multiplication of two matrices written in block form:

$$\mathbf{A} = \begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \qquad \text{and} \qquad \mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{pmatrix} \tag{17}$$

  may be written succinctly as

$$\mathbf{A}\mathbf{B} = \begin{pmatrix} \mathbf{A}_{11}\mathbf{B}_{11} + \mathbf{A}_{12}\mathbf{B}_{21} & \mathbf{A}_{11}\mathbf{B}_{12} + \mathbf{A}_{12}\mathbf{B}_{22} \\ \mathbf{A}_{21}\mathbf{B}_{11} + \mathbf{A}_{22}\mathbf{B}_{21} & \mathbf{A}_{21}\mathbf{B}_{12} + \mathbf{A}_{22}\mathbf{B}_{22} \end{pmatrix} \tag{18}$$

  provided the necessary dimensions work out.

Matrix Operations — Trace, Determinants, Cofactors

These operations are all applied to *square* matrices, i.e. matrices of dimension $n \times n$ for integer $n \geq 1$:

- The *trace* of a matrix $\mathbf{A}$ is the sum of the elements on its main diagonal.

- The *determinant* of a matrix, $\mathbf{A}$, is written

$$\det(\mathbf{A}) \qquad \text{or} \qquad |\,\mathbf{A}\,|\,. \tag{19}$$

  Although this is not the way it is commonly defined in linear algebra texts, the definition I prefer to give is recursive:

  - As a preliminary, let $\mathbf{M}_{ij}$ be the $(n-1) \times (n-1)$ matrix obtained from $\mathbf{A}$ by deleting row i and column j.
  - When $n = 1$,

$$\det(\mathbf{A}) = a_{11}. \tag{20}$$

  - When $n = 2$,

$$\det(\mathbf{A}) = a_{11}a_{22} - a_{12}a_{21}. \tag{21}$$

  - For $n \geq 2$, choose any row i of $\mathbf{A}$. Then,

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{M}_{ij}). \tag{22}$$

    Alternatively, we can choose any column j of $\mathbf{A}$:

$$\det(\mathbf{A}) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(\mathbf{M}_{ij}). \tag{23}$$

- Properties of the determinant:
  - The determinant changes sign when two rows (or columns) are interchanged; it is 0 when two rows (or two columns) are repeated.
  - $\det(\mathbf{I}) = 1$.
  - $\det(\mathbf{AB}) = \det(\mathbf{A})\ \det(\mathbf{B})$.
  - $\det(\mathbf{A}^\top) = \det(\mathbf{A})$.
- The *cofactor matrix* of $\mathbf{A}$ is an $n \times n$ matrix defined as

$$\mathbf{A}^* = \big((-1)^{i+j} \det(\mathbf{M}_{ij})\big). \tag{24}$$

  The determinant $\det(\mathbf{M}_{ij})$ is called the *minor* of a_{ij}. The transpose of the cofactor matrix is called the *adjoint matrix*.

Matrix Operations — Inverse

- For matrix $\mathbf{A}$, a matrix $\mathbf{B}$ such that

$$\mathbf{AB} = \mathbf{I} \tag{25}$$

  is said to be the *right inverse* of $\mathbf{A}$. A matrix $\mathbf{C}$ such that

$$\mathbf{CA} = \mathbf{I} \tag{26}$$

  is said to be the *left inverse* of $\mathbf{A}$.
- For a square matrix $\mathbf{A}$, the left inverse exists if and only if the right inverse exists. In this case, the two inverses are equal and $\mathbf{A}$ is said to be *invertible*. The inverse is generally denoted

$$\mathbf{A}^{-1}. \tag{27}$$

- A matrix that is not invertible is called *singular*.

- Important properties:
  - $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.
  - If $\mathbf{A}$ and $\mathbf{B}$ are both $n \times n$ and both invertible, then

$$(\mathbf{AB})^{-1} = \mathbf{B}^{-1}\mathbf{A}^{-1}. \tag{28}$$

  - $\det(\mathbf{A}^{-1}) = 1/\det(\mathbf{A})$.
  - The inverse depends on the determinant and the adjoint matrix:

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{A}^{*\top} \tag{29}$$

Special Forms of Square Matrices

- A *diagonal matrix* $\mathbf{A}$ is a square matrix where $a_{ij} = 0$ if $i \neq j$. Diagonal matrices are often written using the notation:

$$\operatorname{diag}(a_{11}, \ldots, a_{nn}) \qquad \text{or} \qquad \operatorname{diag}(a_1, \ldots, a_n) \tag{30}$$

- A *lower triangular matrix* $\mathbf{A}$ has $a_{ij} = 0$ whenever $i < j$.
- An *upper triangular matrix* $\mathbf{A}$ has $a_{ij} = 0$ whenever $i > j$.
- Products of diagonal matrices are diagonal. Products of upper triangular matrices are upper triangular. Products of lower triangular matrices are lower triangular.
- Determinant calculations are trivial for these special forms:

$$\det \mathbf{A} = \prod_{i=1}^{n} a_{ii}. \tag{31}$$

A Special Note

When faced with learning about a new mathematical technique, especially in a relatively short amount of time, how should you approach it? You should play! Play with small examples of each idea to try to cement in your mind what it really means. Guess new properties and explore relationships between properties. Start small and build up. You can't address major problems without a firm understanding of smaller ones.

The discussion in this lecture has moved fairly quickly through introductory material on vectors and matrices. Most of it is assumed to be review because it is the initial material taught it almost any linear algebra course. Use small examples to review what you may have forgotten and to learn (develop a "feel for") things that are new.

Potential Test Questions

1. Form two 3×2 matrices, $\mathbf{A}$ and $\mathbf{B}$. Compute $\mathbf{A} + \mathbf{B}$, $\mathbf{C} = \mathbf{A}\mathbf{B}^\top$, and $\mathbf{D} = \mathbf{A}^\top\mathbf{B}$. Compute the trace and determinants of $\mathbf{C}$ and $\mathbf{D}$. Which of these two has an inverse? Calculate this inverse.

2. For what special matrix structures is $\mathbf{AB} = \mathbf{BA}$?

3. Construct two matrices $\mathbf{A}$ and $\mathbf{B}$, such that $\mathbf{AB} = \mathbf{I}$, but $\mathbf{BA} \neq \mathbf{I}$.

4. Let $\mathbf{u} = (u_1, u_2, u_3)^\top$ and $\mathbf{v} = (v_1, v_2, v_3)^\top$. Convert $\mathbf{u}$ to a matrix, $\mathbf{U}$, such that
$$\mathbf{U}\mathbf{v} = \mathbf{u} \times \mathbf{v}.$$

5. (This one is a bit harder.) Matrices can be written in a block form that identifies submatrices, row vectors, and column vectors. For example, an $n \times n$ symmetric, invertible matrix $\mathbf{A}$ can be written
$$\mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{v} \\ \mathbf{v}^\top & c \end{pmatrix},$$
for $(n-1) \times (n-1)$ invertible, symmetric matrix $\mathbf{B}$, $(n-1) \times 1$ column vector, $\mathbf{v}$, and scalar c. Find the inverse of $\mathbf{A}$ as a function of $\mathbf{B}$, $\mathbf{v}$ and c. Hint: handle the possibilities of $c = 0$ and $c \neq 0$ separately.

Homework 1 Questions

These are due Thursday in class along with the questions from Lecture 1.

1. (**10 points**) Given upper triangular matrices $\mathbf{U}_1$ and $\mathbf{U}_2$ and lower triangular matrices $\mathbf{L}_1$ and $\mathbf{L}_2$. Which of the products $\mathbf{U}_1\mathbf{U}_2$, $\mathbf{L}_1\mathbf{L}_2$, $\mathbf{U}_1\mathbf{L}_1$ or $\mathbf{L}_1\mathbf{U}_1$, if any, have triangular forms? Justify your answer.

   Solution As examples easily show, $\mathbf{U}_1\mathbf{L}_1$ and $\mathbf{L}_1\mathbf{U}_1$ do not have triangular forms. The others do. To show this, consider an entry of $\mathbf{L}_1\mathbf{L}_2$ where $i > j$. All we need to do is show that this is 0. Write out the computation as

   $$\sum_{k=1}^{N} l_{i,k} l'_{k,j}$$

   In the terms involving $k < i$, $l_{i,k} = 0$ so the factor is 0. In the terms involving $k \geq i$, we have $k \geq i > j$ so $k > j$. For these, $l'_{k,j} = 0$. Hence the resulting term is 0.

2. (**10 points**) Matrices are often written in special forms, for example as a stack of row vectors

   $$\mathbf{A} = \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \cdots \\ \mathbf{a}_n^\top \end{pmatrix},$$

   or a sequence of column vectors

   $$\mathbf{B} = \begin{pmatrix} \mathbf{b}_1, & \mathbf{b}_2, & \cdots & \mathbf{b}_n \end{pmatrix}.$$

   Assuming that $\mathbf{A}$ is $n \times m$ and $\mathbf{B}$ is $m \times n$, write expressions for

   $$\mathbf{AB}$$

   and

   $$\mathbf{BA}$$

   in terms of $\mathbf{a}_i$ and $\mathbf{b}_j$.

   Solution: Each entry in $\mathbf{AB}$ is formed by a dot product of one pair of vectors. There are n^2 such dot products. The i-j entry, using the notation above, is

   $$(\mathbf{a}_i^\top \mathbf{b}_j).$$

   The entries in $\mathbf{BA}$ are a bit more difficult to see. First, in

   $$\mathbf{BA} = \begin{pmatrix} \mathbf{b}_1, & \mathbf{b}_2, & \cdots & \mathbf{b}_n \end{pmatrix} \begin{pmatrix} \mathbf{a}_1^\top \\ \mathbf{a}_2^\top \\ \cdots \\ \mathbf{a}_n^\top \end{pmatrix},$$

note that the entries of column $\mathbf{b}_i$ are only multiplied by entries in row $\mathbf{a}_i^\top$. Second, each multiplication involving column $\mathbf{b}_i$ in $\mathbf{B}$ and row $\mathbf{a}_i^\top$ of $\mathbf{A}$ contributes to exactly one entry in the final product. Third, a contribution is made by this pair to each entry in the final product. Together, these lead to the following result:

$$\mathbf{BA} = \sum_i \mathbf{b}_i \mathbf{a}_i^\top .$$