Queens College, CUNY, Department of Computer Science

Numerical Methods

CSCI 361 / 761

Spring 2018

Instructor: Dr. Sateesh Mane

November 22, 2018

8 Lecture 8

8.1 Matrices & linear systems of equations

- We shall study the solution of linear systems of equations by using matrices.
- The topic is also called (applied) **linear algebra.**
- In this lecture we shall briefly review the solution of linear systems of equations.
- We shall study the connection of linear systems of equations to matrices.
- This motivates the study of matrices as mathematical objects in their own right.
- We shall examine some basic definitions and properties of matrices.

8.2 Linear systems of equations, Part 1

- Suppose a and b are constants and x is a variable. Consider the following linear equation:

$$ax = b\,. \tag{8.2.1}$$

- The solution is obviously $x = b/a$. For later use it is better to write $x = a^{-1}b$.
- *Is the solution really obvious?*
  1. If $a = 0$, the above solution is not well defined.
  2. Hence we require $a \neq 0$ as a precondition for a solution to exist.
  3. If $a \neq 0$, the above solution is not only well defined but it is also unique.
  4. *We do not require any conditions on the value of b.*
- Next we consider linear equations involving two variables x and y.
  1. Consider the following linear equation:

$$x + 2y = 3\,. \tag{8.2.2}$$

  2. This is the equation of a straight line in the two-dimensional (x, y) plane.
  3. There are infinitely many solutions, e.g. $(x, y) = (3, 0)$ or $(0, \frac{3}{2})$ or $(1, 1)$, etc.
  4. We say the set of equations is **underdetermined.**
  5. There are not enough equations to fix a unique solution.
  6. To obtain a unique solution for x and y, we require *two* equations.
  7. The unique solution is given by the point where the two lines intersect.
  8. *However, if the straight lines are parallel, then there is no solution.* For example:

$$x + 2y = 3\,, \tag{8.2.3a}$$
$$x + 2y = 4\,. \tag{8.2.3b}$$

- Hence to deal with multiple variables, we must formulate the problem more systematically.
- **See next page.**

8.3 Linear systems of equations, Part 2

- To generalize to multiple variables, we denote the unknowns by $x_1, x_2, \dots, x_n$, where $n > 1$.
- Let us write a set of two linear equations in two variables x_1 and x_2 as follows.

$$a_{11}x_1 + a_{12}x_2 = b_1\,, \tag{8.3.1a}$$

$$a_{21}x_1 + a_{22}x_2 = b_2\,. \tag{8.3.1b}$$

- One elegant technique to solve these equations is to employ **matrices.**
- We express the above equations as follows.

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}\,. \tag{8.3.2}$$

- **Hence we have transformed the problem to the solution of a matrix equation.**
- *This lecture is only a refresher. You must learn the definition of a matrix and the rules for matrix addition and multiplication from a textbook.*
- Let us digress slightly: **what if there are *three equations* for x_1 and x_2, not two?**
  1. For example the equations yield three straight lines, which form the edges of a triangle.
  2. Then there is no solution for x_1 and x_2 which satisfies all three equations simultaneously.
  3. We say the system of equations is **overdetermined.**
  4. To obtain a unique solution, *all three lines must meet at one point.*
  5. Issues like this can arise with systems of equations involving multiple variables.
- For three variables x_1, x_2 and x_3 we write a set of three equation as follows.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}\,. \tag{8.3.3}$$

- For arbitrary $n > 1$, we employ a compressed notation to generalize eq. (8.2.1) and write the equations symbolically as follows:

$$A\,\boldsymbol{x} = \boldsymbol{b}\,. \tag{8.3.4}$$

- The solution is $\boldsymbol{x} = A^{-1}\boldsymbol{b}$.
  1. This is the generalization of the solution $x = a^{-1}b$ for one variable.
  2. Here A^{-1} is the **matrix inverse** of the matrix A.
  3. The generalization of the condition $a \neq 0$ is $\det(A) \neq 0$.
  4. Here $\det(A)$ is the **determinant** of the matrix A.
  5. If $\det(A) \neq 0$ then eq. (8.3.4) has a solution and it is unique.
- In general it is computationally expensive and inefficient to compute the matrix inverse A^{-1}.
- There are computationally better algorithms to solve eq. (8.3.4).
- We shall study some of those algorithms in later lectures.

8.4 Matrices: general remarks

- A **matrix** with n rows and k columns is called an $n \times k$ matrix.
  For example for $n = 2$ and $k = 3$ then

$$M_{2\times 3} = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{bmatrix} . \tag{8.4.1}$$

- The notation m_{ij} will be employed to denote the elements of a matrix M.

- A **square matrix** ($n \times n$ square matrix) has n rows and n columns.
  For example for $n = 2$ then

$$N_{2\times 2} = \begin{pmatrix} n_{11} & n_{12} \\ n_{21} & n_{22} \end{pmatrix} . \tag{8.4.2}$$

- Both types of rectangular brackets or parentheses will be employed to describe a matrix.

- A **column vector** (with n elements) is an $n \times 1$ matrix. For example for $n = 3$ then

$$\boldsymbol{v} = \begin{bmatrix} a \\ b \\ c \end{bmatrix} . \tag{8.4.3}$$

- A **row vector** (with n elements) is a $1 \times n$ matrix. For example for $n = 4$ then

$$\boldsymbol{r} = \begin{bmatrix} \alpha & \beta & \gamma & \delta \end{bmatrix} . \tag{8.4.4}$$

- If A is an $m \times n$ matrix, then we can add (or subtract) a matrix B to it if and only if B is also an $m \times n$ matrix. The two matrices must have the same size.

- If we multiply a matrix M by a number λ, the matrix elements of λM (also $M\lambda$, because $\lambda M = M\lambda$) are obtained by multiplying all the elements of M by λ.
  An example for $n = 2$ and $k = 3$ is (see eq. (8.4.1)

$$\lambda\, M = M\, \lambda = \begin{bmatrix} \lambda m_{11} & \lambda m_{12} & \lambda m_{13} \\ \lambda m_{21} & \lambda m_{22} & \lambda m_{23} \end{bmatrix} . \tag{8.4.5}$$

- Matrix multiplication: if A is an $m \times n$ matrix and B is an $n \times k$ matrix, then the matrix product AB is an $m \times k$ matrix.

  1. **Note that the matrix product BA does not exist if $k \neq m$.**
  2. The order of matrix multiplication is important: AB might exist but BA might not.
  3. Even if both matrix products AB and BA exist, **they may not be equal**:

$$AB \neq BA \qquad\qquad \text{(in general)}\,. \tag{8.4.6}$$

  4. We say that matrix multiplication is **non-commutative.**

8.5 Matrices: more definitions

- The **zero matrix** has all entries equal to zero. Obvious enough.
- For a square matrix, the **leading diagonal** is the one with elements m_{ij} where $i = j$. The leading diagonal is also called the **main diagonal**.
- A **diagonal matrix** has nonzero elements only along its leading diagonal. For $n = 3$ an example is

$$D_{3\times 3} = \begin{pmatrix} d_{11} & 0 & 0 \\ 0 & d_{22} & 0 \\ 0 & 0 & d_{33} \end{pmatrix} . \tag{8.5.1}$$

- The **unit (or identity) matrix** is a diagonal matrix will all nonzero elements equal to 1. The identity matrix is frequently denoted by the symbol I or $\boldsymbol{I}$. The $n \times n$ unit or identity matrix is

$$I_{n\times n} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \ddots & \\ 0 & 0 & \dots & 1 \end{pmatrix} . \tag{8.5.2}$$

- A **tridiagonal matrix** has nonzero entries only on the leading diagonal and the two diagonals immediately above and below the leading diagonal. The matrix must be a square matrix. For $n = 5$ an example is

$$M_{\rm tri} = \begin{pmatrix} a_1 & b_1 & 0 & 0 & 0 \\ c_2 & a_2 & b_2 & 0 & 0 \\ 0 & c_3 & a_3 & b_3 & 0 \\ 0 & 0 & c_4 & a_4 & b_4 \\ 0 & 0 & 0 & c_5 & a_5 \end{pmatrix} . \tag{8.5.3}$$

- A **pentadiagonal matrix** is ... *you figure it out.*
- A **banded matrix** has nonzero elements only in diagonals such that $m_{ij} = 0$ if $|i-j| > c$, for some integer constant $c \geq 0$. A diagonal matrix is a banded matrix with $c = 0$. A tridiagonal matrix is a banded matrix with $c = 1$.
- An **upper triangular matrix** U has nonzero entries only on or above the leading diagonal. A **lower triangular matrix** L has nonzero entries only on or below the leading diagonal. Both the matrices L and U must be square. For $n = 4$, examples are

$$L = \begin{pmatrix} \ell_{11} & 0 & 0 & 0 \\ \ell_{21} & \ell_{22} & 0 & 0 \\ \ell_{31} & \ell_{32} & \ell_{33} & 0 \\ \ell_{41} & \ell_{42} & \ell_{43} & \ell_{44} \end{pmatrix} , \qquad U = \begin{pmatrix} u_{11} & u_{12} & u_{13} & u_{14} \\ 0 & u_{22} & u_{23} & u_{24} \\ 0 & 0 & u_{33} & u_{34} \\ 0 & 0 & 0 & u_{44} \end{pmatrix} . \tag{8.5.4}$$

  For $n \times n$ matrices, the number of independent matrix elements in both cases is $n(n+1)/2$. A diagonal matrix is simultaneously both upper and lower triangular.

8.6 Matrix transpose

- The **transpose** of a matrix M is denoted by M^T.
  The transpose M^T is obtained by interchanging the rows and columns of the matrix M.

- If M is an $n \times k$ matrix then M^T is a $k \times n$ matrix.

- The matrix M **need not be square**.

- The transpose matrix always exists and is unique.

- For $n = 2$ and $k = 3$ an example is

$$M = \begin{bmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \end{bmatrix}, \qquad M^T = \begin{bmatrix} m_{11} & m_{21} \\ m_{12} & m_{22} \\ m_{13} & m_{23} \end{bmatrix}. \tag{8.6.1}$$

- The transpose of a column vector is a row vector and vice versa.

- The transpose of a transpose is the original matrix: $(M^T)^T = M$.

- The product of a matrix and its transpose is a square matrix:
  MM^T is an $n \times n$ square matrix and $M^T M$ is a $k \times k$ square matrix.

- The transpose of the matrix product AB is given by

$$(AB)^T = B^T A^T. \tag{8.6.2}$$

  Similarly

$$(ABC)^T = C^T B^T A^T. \tag{8.6.3}$$

  There is an obvious pattern for products of more matrices.

8.7 Matrix inverse

- The **inverse** of a matrix M is denoted by M^{-1}.
- The matrix M must be square.
- **Not every matrix has an inverse: M^{-1} does not always exist.**
- A matrix which does not have an inverse is called **singular**.
- The inverse matrix M^{-1} (if it exists) is defined by the relations

$$MM^{-1} = I\,, \qquad M^{-1}M = I\,. \tag{8.7.1}$$

- **If the inverse matrix M^{-1} exists, then it is unique.**
- The zero matrix (all entries are zero) is obviously singular.
- The inverse of the unit matrix is obviously the unit matrix itself.
- A matrix can equal its own inverse, i.e. $M^2 = I$. Some examples for $n = 2$ are

$$M = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}\,, \qquad M = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}\,. \tag{8.7.2}$$

- Examples of singular matrices are

$$M_{\text{singular}} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}\,, \qquad M_{\text{singular}} = \begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\,. \tag{8.7.3}$$

- The inverse of the matrix product AB is given by

$$(AB)^{-1} = B^{-1}A^{-1}\,. \tag{8.7.4}$$

  Similarly

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}\,. \tag{8.7.5}$$

  There is an obvious pattern for products of more matrices.

8.7.1 Non-square matrices

- **For your information only. Not for examination.**
- If the matrix M is not square, one can define a "left inverse" L_{inv} and a "right inverse" R_{inv}.
- The theory becomes complicated. Suppose M is $n \times k$, where $n \neq k$.
- If $n > k$, only a left inverse can exist. If it does, there are infinitely many choices. The inverse is not unique.

$$L_{\text{inv}}M = I_{k\times k}\,. \tag{8.7.6}$$

- If $n < k$, only a right inverse can exist. If it does, there are infinitely many choices. The inverse is not unique.

$$MR_{\text{inv}} = I_{n\times n}\,. \tag{8.7.7}$$

8.8 Square matrices: commutation & anticommutation

- Square matrices have some special properties which are important in many applications.
- Consider two square matrices A and B, both of the same size $(n \times n)$.
- Then A and B are said to **commute** if $AB = BA$.
- Also A and B are said to **anticommute** if $AB = -BA$.

8.9 Square matrices: symmetric & antisymmetric

- A **symmetric matrix** S is equal to its transpose: $S^T = S$.
- An **antisymmetric matrix** A is equal to the **negative** of its transpose: $A^T = -A$.
- An antisymmetric matrix is also called a **skew-symmetric matrix**.
- Hence the matrix elements satisfy the relations

$$(S^T)_{ij} = S_{ij}\,, \qquad (A^T)_{ij} = -A_{ij}\,. \tag{8.9.1}$$

- The elements on the leading diagonal of an antisymmetric matrix are all zero.
- Every diagonal matrix is symmetric.
- For $n = 3$ an example is

$$S_{3\times 3} = \begin{pmatrix} a_{11} & b & c \\ b & a_{22} & d \\ c & d & a_{33} \end{pmatrix}\,, \qquad A_{3\times 3} = \begin{pmatrix} 0 & \beta & \gamma \\ -\beta & 0 & \delta \\ -\gamma & -\delta & 0 \end{pmatrix}\,. \tag{8.9.2}$$

- An $n \times n$ symmetric matrix has $n(n+1)/2$ independent elements.
- An $n \times n$ antisymmetric matrix has $n(n-1)/2$ independent elements.
- The sum of two symmetric matrices is also symmetric.
- The sum of two antisymmetric matrices is also antisymmetric.
- The product of a symmetric matrix with a number is a symmetric matrix.
- The product of an antiymmetric matrix with a number is an antisymmetric matrix.
- **However, do not jump to conclusions when we multiply symmetric and antisymmetric matrices.**
- For any arbitrary matrix M, the two products MM^T and M^TM are both square symmetric matrices. Let M be an $n \times k$ matrix. It was pointed out in Sec. 8.6 that MM^T is an $n \times n$ square matrix and M^TM is a $k \times k$ square matrix.

  1. **Proof that MM^T is symmetric.**
     Let $S_1 = MM^T$. Then $S_1^T = (MM^T)^T = (M^T)^TM^T = MM^T = S_1$.
     Hence $S_1^T = S_1$ so MM^T is symmetric.
  2. **Proof that M^TM is symmetric.**
     Let $S_2 = M^TM$. Then $S_2^T = (M^TM)^T = M^T(M^T)^T = M^TM = S_2$.
     Hence $S_2^T = S_2$ so M^TM is symmetric.

8.9.1 Products of symmetric and antisymmetric matrices

- **Do not jump to conclusions about the products of symmetric and antisymmetric matrices.**
- The following matrices are symmetric:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} . \tag{8.9.3}$$

  (Never mind why I call them σ_1 and σ_3. Yes there is also a σ_2.)

- The matrix products $\sigma_3\sigma_3$ and $\sigma_1\sigma_1$ are also symmetric matrices ($\sigma_3^2 = \sigma_1^2 = I$):

$$\sigma_3\sigma_3 = \sigma_1\sigma_1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} . \tag{8.9.4}$$

- **The matrix products $\sigma_1\sigma_3$ and $\sigma_3\sigma_1$ are antisymmetric:**

$$\sigma_1\sigma_3 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} , \qquad \sigma_3\sigma_1 = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} . \tag{8.9.5}$$

- The symmetric matrices σ_1 and σ_3 anticommute. Their product is an antisymmetric matrix.
- Yes the product is proportional to σ_2. If you really want to know,

$$\sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} . \tag{8.9.6}$$

  The matrix elements of σ_2 are complex numbers (pure imaginary).

8.10 Square matrices: trace

- The **trace** of a square matrix is the sum of the elements on the main diagonal.
- Some authors also refer to the trace as the **spur** of a matrix.

$$\mathrm{spur}(M) = \mathrm{trace}(M) = \mathrm{Tr}(M) = m_{11} + m_{22} + \cdots + m_{nn} = \sum_{i=1}^{n} m_{ii}\,. \qquad (8.10.1)$$

- A **traceless** matrix is one whose trace equals zero.
- An antisymmetric matrix is always traceless.
- The trace of a square matrix is equal to the trace of its transpose:

$$\mathrm{Tr}(M) = \mathrm{Tr}(M^T)\,. \qquad (8.10.2)$$

- For two square matrices A and B, the trace of $A+B$ is the sum of the traces

$$\mathrm{Tr}(A+B) = \mathrm{Tr}(A) + \mathrm{Tr}(B)\,. \qquad (8.10.3)$$

- Furthermore, for matrices A and B,

$$\mathrm{Tr}(AB) = \mathrm{Tr}(BA)\,. \qquad (8.10.4)$$

- The trace of a matrix product has a cyclic property. For three matrices A, B and C,

$$\mathrm{Tr}(ABC) = \mathrm{Tr}(BCA) + \mathrm{Tr}(CAB)\,. \qquad (8.10.5)$$

- The cyclic property extends in an obvious way to products of more matrices.

8.11 Square matrices: determinant

- The **determinant** of a square matrix is a number obtained from all the rows and colums of the matrix. For a 2×2 matrix, the determinant is given by

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \qquad \det(M) = ad - bc\,. \tag{8.11.1}$$

- In general, for an $n \times n$ square matrix, the determinant is given by a sum of products

$$\det(M) = \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \cdots \sum_{i_n=1}^{n} \sigma_{i_1,i_2,\dots,i_n} m_{1i_1} m_{2i_2} \dots m_{ni_n}\,. \tag{8.11.2}$$

  Here $\sigma_{i_1,i_2,\dots,i_n}$ is a sign factor equal to ± 1, given by

$$\sigma_{i_1,i_2,\dots,i_n} = \begin{cases} +1 & ((i_1, i_2, \dots, i_n) = \text{even permutation of } (1,2,\dots,n)) \\ \\ -1 & ((i_1, i_2, \dots, i_n) = \text{odd permutation of } (1,2,\dots,n))\,. \end{cases} \tag{8.11.3}$$

- Hence in every product in eq. (8.11.2), **each row i and each row j appears exactly once.**
- There are computationally more efficient ways to calculate the determiant of a matrix.
- The determinant of a unit matrix equals unity: $\det(I) = 1$.
- The determinant of a matrix is equal to the determinant of its transpose:

$$\det(M) = \det(M^T)\,. \tag{8.11.4}$$

- For matrices $A_1, \dots, A_n$ the determinant of the product equals the product of the determinants

$$\det(A_1 A_2 \dots A_n) = \det(A_1)\, \det(A_2)\, \dots \det(A_n)\,. \tag{8.11.5}$$

- The inverse M^{-1} of a square matrix M exists if and only if $\det(M) \neq 0$.
  1. By definition, $\det(MM^{-1}) = \det(I) = 1$.
  2. Hence $\det(MM^{-1}) = \det(M)\det(M^{-1}) = 1$.
  3. Therefore $\det(M)$ cannot equal zero.
  4. It is more difficult to prove that if $\det(M) \neq 0$ then the inverse matrix exists.
- **It is possible for the product of two nonzero matrices to be zero.**
  1. It is possible that $AB = 0$ even though $A \neq 0$ and $B \neq 0$.
  2. Then $\det(AB) = \det(A)\det(B) = 0$.
  3. Hence either $\det(A) = 0$ or $\det(B) = 0$ or both.

8.12 Determinant and minors

- Let M be an $n \times n$ square matrix with elements m_{ij}, $i,j = 1,\dots,n$.
- A common notation for the determinant $\det(M)$ is the following

$$\det(M) = \begin{vmatrix} m_{11} & m_{12} & \dots & m_{1n} \\ m_{21} & m_{22} & \dots & m_{2n} \\ \vdots & \vdots & & \vdots \\ m_{n1} & m_{n2} & \dots & m_{nn} \end{vmatrix} . \tag{8.12.1}$$

- We next define the **minor** A_{ij} of the element m_{ij} of M.
- The minor A_{ij} is the determinant of the $(n-1)\times(n-1)$ square matrix obtained by excluding row i and column j from M.
- Hence the matrix M has n^2 minors, one for each element m_{ij}.
- Consider the example of $n = 3$. Then

$$M = \begin{pmatrix} m_{11} & m_{12} & m_{13} \\ m_{21} & m_{22} & m_{23} \\ m_{31} & m_{32} & m_{33} \end{pmatrix} . \tag{8.12.2}$$

- The minors of M are

$$\begin{aligned} A_{11} &= \begin{vmatrix} m_{22} & m_{23} \\ m_{32} & m_{33} \end{vmatrix} & A_{12} &= \begin{vmatrix} m_{21} & m_{23} \\ m_{31} & m_{33} \end{vmatrix} & A_{13} &= \begin{vmatrix} m_{21} & m_{22} \\ m_{31} & m_{32} \end{vmatrix} \\ A_{21} &= \begin{vmatrix} m_{12} & m_{13} \\ m_{32} & m_{33} \end{vmatrix} & A_{22} &= \begin{vmatrix} m_{11} & m_{13} \\ m_{31} & m_{33} \end{vmatrix} & A_{23} &= \begin{vmatrix} m_{11} & m_{12} \\ m_{31} & m_{32} \end{vmatrix} \\ A_{31} &= \begin{vmatrix} m_{12} & m_{13} \\ m_{22} & m_{23} \end{vmatrix} & A_{32} &= \begin{vmatrix} m_{11} & m_{13} \\ m_{21} & m_{23} \end{vmatrix} & A_{33} &= \begin{vmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{vmatrix} . \end{aligned} \tag{8.12.3}$$

- The determinant $\det(M)$ is given by the recursive formula

$$\det(M) = m_{11}A_{11} - m_{12}A_{12} + \cdots + (-1)^{j-1} m_{ij}A_{ij} + \cdots + (-1)^{n-1} m_{in}A_{in} . \tag{8.12.4}$$

- This is clearly a computationally inefficient procedure.
- Gaussian elimination (see Sec. 8.14 below) is much more efficient.

8.13 Matrix inverse: Cramer's rule

- **Cramer's rule** is an explicit formula to calculate the inverse of a matrix.
- However, Cramer's rule is computationally inefficient and not suitable for most applications.
- Let M be an $n \times n$ square matrix with elements m_{ij}, $i, j = 1, \dots, n$.
- Then the inverse matrix M^{-1} is given by using the minors:

$$(M^{-1})_{ij} = (-1)^{i+j} \frac{A_{ji}}{\det(M)} . \tag{8.13.1}$$

- Writing it out explicitly, we obtain

$$M^{-1} = \frac{1}{\det(M)} \begin{pmatrix} A_{11} & -A_{21} & \dots & (-1)^{1+n} A_{n1} \\ -A_{12} & A_{22} & \dots & (-1)^{2+n} A_{n2} \\ \vdots & \vdots & & \vdots \\ (-1)^{n+1} A_{1n} & (-1)^{n+2} A_{2n} & \dots & A_{nn} \end{pmatrix} . \tag{8.13.2}$$

- Cramer's rule yields the inverse for a 2×2 matrix:

$$\begin{aligned} M &= \begin{pmatrix} a & b \\ c & d \end{pmatrix} , \\ M^{-1} &= \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} . \end{aligned} \tag{8.13.3}$$

8.14 Determinants & Gaussian elimination

- The determinant of a diagonal matrix is the product of its diagonal elements. This is obvious. From eq. (8.11.2), any product which contains a non-diagonal matrix element equals zero. Therefore only the product of the diagonal elements contributes to the determinant.

- The determinant of a lower triangular matrix is the product of its diagonal elements. The proof is similar to that for a diagonal matrix. Recall that in every product in eq. (8.11.2), each row i and each row j appears exactly once. Hence in eq. (8.11.2), any product (except the product of the diagonal elements) *must contain one or more elements above the main diagonal,* therefore it contributes zero. Hence only the product of the diagonal elements contributes to the determinant.

- The determinant of an upper triangular matrix is the product of its diagonal elements. The proof is similar to that for a lower triangular matrix.

- **For any square matrix, if we add a multiple of one row to another row, the value of the determinant does not change.**

- **For any square matrix, if we add a multiple of one column to another column, the value of the determinant does not change.**

- The above facts can be used to add/subtract multiples of the rows (or columns) of a matrix from each other to eliminate all the elements below the main diagonal, to obtain an upper triangular matrix. The determinant of the original matrix equals the determinant of the upper triangular matrix (the product of the diagonal elements of the upper triangular matrix).

- One can perform the elimination to obtain a lower triangular matrix. The determinant of the original matrix equals the product of the diagonal elements of the lower triangular matrix.

- **Gaussian elimination** is the process of adding multiples of rows of a matrix to each other to obtain an upper (or lower) triangular matrix.

- **Gauss–Jordan elimination** is the process of adding multiples of rows of a matrix to each other to obtain a diagonal matrix.

8.15 Determinants & swapping rows and/or columns

- **If we swap two rows of a matrix, the determinant changes sign.**
- **If we swap two columns of a matrix, the determinant changes sign.**
- To illustrate, consider the simple example of a 3×3 diagonal matrix:

$$M = \begin{pmatrix} a & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & c \end{pmatrix} . \tag{8.15.1}$$

- The determinant is $\det(M) = abc$.
- Swap rows 1 and 2 and we obtain

$$M_1(\text{swap rows 1 and 2}) = \begin{pmatrix} 0 & b & 0 \\ a & 0 & 0 \\ 0 & 0 & c \end{pmatrix} . \tag{8.15.2}$$

- The determinant is (using Cramer's rule) $\det(M_1) = -bA_{12} = -bac = -abc = -\det(M)$.
- Swap columns 1 and 3 and we obtain

$$M_2(\text{swap columns 1 and 3}) = \begin{pmatrix} 0 & 0 & a \\ 0 & b & 0 \\ c & 0 & 0 \end{pmatrix} . \tag{8.15.3}$$

- The determinant is (using Cramer's rule) $\det(M_2) = cA_{13} = a(-bc) = -abc = -\det(M)$.
- **The combined ideas of swapping rows and Gaussian elimination yields computationally efficient algorithms to solve matrix equations.**
- **The combined ideas of swapping rows and columns and Gauss–Jordan elimination also yields computationally efficient algorithms to solve matrix equations.**

8.16 Matrix equations

- Suppose M is an $n \times n$ square matrix and $\boldsymbol{v}$ is an n–component column vector.
- Suppose also that $\boldsymbol{x}$ is an n–component column vector. We wish to solve the matrix equation

$$M\boldsymbol{x} = \boldsymbol{v}\,. \tag{8.16.1}$$

- The formal solution of eq.(8.16.1) is

$$\boldsymbol{x} = M^{-1}\boldsymbol{v}\,. \tag{8.16.2}$$

- In general, eq. (8.16.2) is computationally inefficient because the calculation of the matrix inverse M^{-1} is computationally expensive.
- There are computationally more efficient ways to solve eq.(8.16.1), which employ ideas of Gaussian or Gauss–Jordan elimination (together with swapping of rows and/or columns).
- However in the simple case $n = 2$, eq. (8.16.2) is easy to apply.
- For $n = 2$, the equation and solution are

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\,, \\ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= M^{-1}\boldsymbol{v} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\,. \end{aligned} \tag{8.16.3}$$

- Explicitly, the solution is

$$x_1 = \frac{dv_1 - bv_2}{ad-bc}\,, \qquad x_2 = \frac{av_2 - cv_1}{ad-bc}\,. \tag{8.16.4}$$

8.17 Orthogonal matrices

- **The material in this section is not for examination.**
- **You may skip it if it is too advanced.**
- An **orthogonal matrix** O satisfies the relation

$$O^T O = I\,. \tag{8.17.1}$$

- It is also true that $OO^T = I$ and $O^T = O^{-1}$.
- The determinant of an orthogonal matrix equals ± 1.
- The converse is not true: a square matrix with determinant 1 is not necessarily orthogonal

$$M = \begin{pmatrix} 2 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} \qquad\qquad (\textbf{not orthogonal})\,. \tag{8.17.2}$$

- An orthogonal matrix with determinant $+1$ is called a **special orthogonal matrix**.
- The identity matrix is a special orthogonal matrix.
- The negative of the identity matrix $-I_{n\times n}$ has determinant $(-1)^n$.
  It equals 1 if n is even and -1 if n is odd.
- Examples of symmetric and antisymmetric special orthogonal matrices are
  (x is a real number)

$$S = \begin{pmatrix} \cosh(x) & \sinh(x) \\ \sinh(x) & \cosh(x) \end{pmatrix}\,, \qquad A = \begin{pmatrix} \cos(x) & -\sin(x) \\ \sin(x) & \cos(x) \end{pmatrix}\,. \tag{8.17.3}$$

- The following matrix is orthogonal and has determinant -1

$$O = \begin{pmatrix} \cos(x) & \sin(x) \\ \sin(x) & -\cos(x) \end{pmatrix}\,. \tag{8.17.4}$$

- Rotation matrices, in any number of dimensions, are special orthogonal matrices.
- Reflection matrices are orthogonal matrices, not always special orthogonal matrices.

8.17.1 Invariant

- If O is an orthogonal matrix and $O\boldsymbol{u} = \boldsymbol{v}$, then $\boldsymbol{u}^T\boldsymbol{u} = \boldsymbol{v}^T\boldsymbol{v}$.
- The proof is simple: $\boldsymbol{v}^T\boldsymbol{v} = (\boldsymbol{u}^T O^T)(O\boldsymbol{u}) = \boldsymbol{u}^T(O^T O)\boldsymbol{u} = \boldsymbol{u}^T\boldsymbol{u}$.
- Hence the value of $\boldsymbol{u}^T\boldsymbol{u}$ is **invariant** if $\boldsymbol{u}$ is multiplied by an orthogonal matrix O.

8.18 Complex valued matrices

- **The material in this section is not for examination.**
- **You may skip it if it is too advanced.**
- The elements of a matrix need not be real. They can be complex.
- The **Hermitian conjugate** of a complex valued matrix C is the complex conjugate transpose matrix
$$C^\dagger = (C^T)^* = (C^*)^T \,. \tag{8.18.1}$$
- A **Hermitian matrix** H is equal to its Hermitian conjugate
$$H^\dagger = H \,. \tag{8.18.2}$$
- An **anti-Hermitian** or **skew-Hermitian** matrix A is equal to the negative of its Hermitian conjugate
$$A^\dagger = -A \,. \tag{8.18.3}$$
- The eigenvalues of a Hermitian matrix are real numbers.
The eigenvalues of a skew-Hermitian matrix are pure imaginary numbers.
- A **unitary matrix** U satisfies the relation
$$U^\dagger U = I \,. \tag{8.18.4}$$
- It is also true that $UU^\dagger = I$ and $U^\dagger = U^{-1}$.
- The determinant of a unitary matrix satisfies $|\det U| = 1$.
Hence the determinant of a unitary matrix lies on the complex unit circle.
- The eigenvalues of a unitary matrix lie on the complex unit circle.
- A unitary matrix with determinant $+1$ is called a **special unitary matrix**.
- The unit matrix is both a Hermitian matrix and a special unitary matrix.
- Examples of symmetric and antisymmetric special unitary matrices are
$$S = \begin{pmatrix} 0 & i \\ i & 0 \end{pmatrix} , \qquad A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} . \tag{8.18.5}$$
- The following matrices are unitary and have determinant -1 and $-i$, respectively
$$iI_{2\times 2} = \begin{pmatrix} i & 0 \\ 0 & i \end{pmatrix} , \tag{8.18.6}$$
$$iI_{3\times 3} = \begin{pmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & i \end{pmatrix} . \tag{8.18.7}$$

8.19 Matrix exponential

- **The material in this section is not for examination.**
- **You may skip it if it is too advanced.**
- The **exponential** of a square matrix is **defined** by the power series infinite sum

$$\exp(M) = e^M = 1 + M + \frac{M^2}{2!} + \frac{M^3}{3!} + \cdots = \sum_{n=0}^{\infty} \frac{M^n}{n!} \,. \tag{8.19.1}$$

- Note that this is a **definition.**
- As with the exponential series for a number x, the exponential series for a matrix converges absolutely.
- Be careful: if A and B are square matrices, then in general

$$e^A e^B \neq e^{A+B} \,. \tag{8.19.2}$$

  The above result is true only if A and B commute.
- The exponential of an antisymmetric matrix is a special orthogonal matrix

$$(e^A)^T e^A = e^{A^T} e^A = e^{-A} e^A = I \,. \tag{8.19.3}$$

- The exponential of an anti-Hermitian matrix is a special unitary matrix

$$(e^A)^\dagger e^A = e^{A^\dagger} e^A = e^{-A} e^A = I \,. \tag{8.19.4}$$

8.20 Matrix rank

- **The material in this section is not for examination.**
- **You may skip it if it is too advanced.**
- The matrix M need not be square.
- The **rank** of a matrix M is the dimension of the vector space spanned by the columns of M.
- The rank of M is also equal to the dimension of the vector space spanned by the rows of M.
- Hence the column rank equals the row rank of a matrix.
- **We shall see a little bit of this when we solve a set of simultaneous linear equations, and the equations are not all linearly independent. In that case, if there are n equations, the rank will be less than n. It may or may not be possible to solve the equations in this situation. We shall see.**
- **Nevertheless, concepts such as "vector space" are** not for examination.