Lecture 7

Matrix, Vector Operations (Linear Algebra Review)

Owen L. Lewis

Department of Mathematics and Statistics
University of New Mexico

Sept. 10, 2024

Goals:

- Newton's Method in higher dimension. . . A MATRIX
- recall linear algebra ouch!
- cost analysis of basic operations
- identify solution schemes to systems

Newton's Method in Higher Dimensions

Consider roots of $\mathbf{f}(\mathbf{x}) = \mathbf{0}$ where,

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} f_1(\mathbf{x}) \\ f_2(\mathbf{x}) \\ \vdots \\ f_n(\mathbf{x}) \end{bmatrix}, \quad \mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad \mathbf{0} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Notes

To the board!

The Jacobian

The “high dimensional derivative” of $\mathbf{f}$ is called the “Jacobian” and it is a matrix:

$$\mathbf{J_f}(\mathbf{x}) = \begin{bmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \frac{\partial f_n}{\partial x_2} & \cdots & \frac{\partial f_n}{\partial x_n} \end{bmatrix}.$$

Note: this matrix encodes all of the derivatives of $\mathbf{f}$ with respect to all of the entries of $\mathbf{x}$. Its entries are functions! If we evaluate it at any specific $\mathbf{x}$, it becomes a normal matrix of numbers.

Newton's Method in Higher Dimensions

We usually write

$$\mathbf{x}_{n+1} = \mathbf{x}_n - \left[\mathbf{J}_{\mathbf{F}}(\mathbf{x}_n)\right]^{-1}\mathbf{F}(\mathbf{x}_n)$$

because it has nice analogies to

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

However, the *actual* algorithm solves

$$\mathbf{J}_{\mathbf{F}}(\mathbf{x}_n)\mathbf{h}_n = -\mathbf{F}(\mathbf{x}_n)$$

and then set $\mathbf{x}_{k+1} = \mathbf{x}_k + \mathbf{h}_k$. This is how Newton's Method *really* works.

Example:Newton in 2-D

Example: Consider the problem

$$\sin(x) + y + 4 = 0, \quad x^2 = xy.$$

First, rewrite:

$$\mathbf{f}(\mathbf{x}) = \begin{bmatrix} \sin(x) + y + 4 \\ x^2 - xy \end{bmatrix} = \mathbf{0}.$$

The Jacobian is:

$$\mathbf{J}_{\mathbf{f}}(\mathbf{x}) = \begin{bmatrix} \cos(x) & 1 \\ 2x - y & -x \end{bmatrix}$$

Start with an initial guess: $\mathbf{x}_0 = [1, -2]^T$:

$$\begin{bmatrix} 0.54 & 1 \\ 4 & -1 \end{bmatrix} \mathbf{h}_0 = \begin{bmatrix} -2.8 \\ -3 \end{bmatrix}$$

We get $\mathbf{h}_0 = [-1.2, -2.1]^T$ and so

$$\mathbf{x}_1 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} + \begin{bmatrix} -1.2 \\ -2.1 \end{bmatrix} = \begin{bmatrix} -0.2 \\ -4.1 \end{bmatrix}$$

Matrix Inversion

Solving systems of *linear* equations is necessary for each step of Newton's Method.

Solving systems of linear equations is equivalent to solving a matrix equation.

Solving a matrix equation is equivalent to finding its inverse matrix.

We will learn A LOT about this in the next two or three weeks.

Chapter 2: Solution of Linear Systems

- matrix problems arise in many areas of math, science and engineering
- Basic Linear Algebra Subprograms (BLAS) is an interface standard for operations
- simple systems set the stage for further development: avoiding error, avoiding large costs

Prereq

Linear Algebra used to be a prerequisite of the course!

- you should feel comfortable with Appendix A in your textbook
- We'll revisit some linear algebra operations

Notes

Matrix Notation

The matrix A with m rows and n columns looks like:

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

$a_{ij} =$ element in **row** i, and **column** j

Matrices Consist of Row and Column Vectors

As a collection of column vectors

$$A = \left[\begin{array}{c|c|c|c} & & & \\ a_{(1)} & a_{(2)} & \cdots & a_{(n)} \\ & & & \end{array}\right]$$

As a collection of row vectors

$$A = \left[\begin{array}{c} a'_{(1)} \\ \hline a'_{(2)} \\ \hline \vdots \\ \hline a'_{(m)} \end{array}\right]$$

A prime is used to designate a row vector on this and the following pages.

Matrix–Vector Product

- The Row View
  - easy to do calculations by hand
- The Column View
  - gives mathematical insight

Row View of Matrix–Vector Product

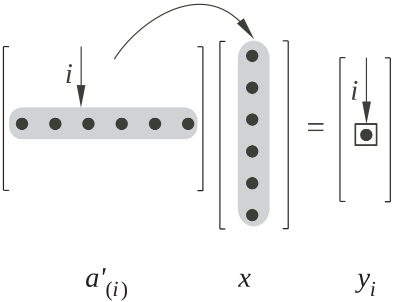
Product of a 3×4 matrix, A, with a 4×1 vector, x, looks like

$$\begin{bmatrix} \vec{a}'_{(1)} \\ \hline \vec{a}'_{(2)} \\ \hline \vec{a}'_{(3)} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} \vec{a}'_{(1)} \cdot \vec{x} \\ \vec{a}'_{(2)} \cdot \vec{x} \\ \vec{a}'_{(3)} \cdot \vec{x} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}$$

where $\vec{a}'_{(1)}$, $\vec{a}'_{(2)}$, and $\vec{a}'_{(3)}$, are the *row vectors* constituting the A matrix.

The matrix–vector product $b = Ax$ produces elements in b by forming inner products of the rows of A with x.

Row View of Matrix–Vector Product



Compatibility Requirement

Inner dimensions must agree

$$\begin{array}{cccc} A & x & = & b \\ [m \times n] & [n \times 1] & = & [m \times 1] \end{array}$$

Row View of Matrix–Vector Product

Consider the following matrix–vector product written out as a dot product with matrix rows

$$\begin{bmatrix} 5 & 0 & 0 & -1 \\ -3 & 4 & -7 & 1 \\ 1 & 2 & 3 & 6 \end{bmatrix} \begin{bmatrix} 4 \\ 2 \\ -3 \\ -1 \end{bmatrix}$$

$$= \begin{bmatrix} (5)(4) & + & (0)(2) & + & (0)(-3) & + & (-1)(-1) \\ (-3)(4) & + & (4)(2) & + & (-7)(-3) & + & (1)(-1) \\ (1)(4) & + & (2)(2) & + & (3)(-3) & + & (6)(-1) \end{bmatrix} = \begin{bmatrix} 21 \\ 16 \\ -7 \end{bmatrix}$$

This is the row view.

Row View of Matrix–Vector Product

$$= \begin{bmatrix} (5)(4) & + & (0)(2) & + & (0)(-3) & + & (-1)(-1) \\ (-3)(4) & + & (4)(2) & + & (-7)(-3) & + & (1)(-1) \\ (1)(4) & + & (2)(2) & + & (3)(-3) & + & (6)(-1) \end{bmatrix}$$

Now, group the multiplication and addition operations by column:

$$4 \begin{bmatrix} 5 \\ -3 \\ 1 \end{bmatrix} + 2 \begin{bmatrix} 0 \\ 4 \\ 2 \end{bmatrix} - 3 \begin{bmatrix} 0 \\ -7 \\ 3 \end{bmatrix} - 1 \begin{bmatrix} -1 \\ 1 \\ 6 \end{bmatrix} = \begin{bmatrix} 21 \\ 16 \\ -7 \end{bmatrix}$$

This is just a combination of the column vectors, with weights given by the vector $\vec{x}$. Final result is identical to that obtained with the row view.

Column View of Matrix–Vector Product

Consider a **linear combination of a set of column vectors** $\{\vec{a}_{(1)}, \vec{a}_{(2)}, \ldots, \vec{a}_{(n)}\}$. Each $\vec{a}_{(j)}$ has m elements
Let x_i be a set (a vector) of scalar multipliers

$$x_1\vec{a}_{(1)} + x_2\vec{a}_{(2)} + \ldots + x_n\vec{a}_{(n)} = \vec{b}$$

or

$$\sum_{j=1}^{n} \vec{a}_{(j)} x_j = \vec{b}$$

Expand the (hidden) row index

$$x_1 \begin{bmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{m1} \end{bmatrix} + x_2 \begin{bmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{m2} \end{bmatrix} + \cdots + x_n \begin{bmatrix} a_{1n} \\ a_{2n} \\ \vdots \\ a_{mn} \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Column View of Matrix–Vector Product

Form a matrix with the $a_{(j)}$ as columns

$$\begin{bmatrix} \Bigg| & \Bigg| & & \Bigg| \\ \vec{a}_{(1)} & \vec{a}_{(2)} & \cdots & \vec{a}_{(n)} \\ \Bigg| & \Bigg| & & \Bigg| \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$

Or, writing out the elements

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Column View of Matrix–Vector Product

Thus, the matrix-vector product is

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

Save space with matrix notation

$$Ax = b$$

Column View of Matrix–Vector Product

The matrix–vector product $b = Ax$ produces a vector b from a linear combination of the columns in A.

$$\vec{b} = \vec{A}x \quad \Longleftrightarrow \quad b_i = \sum_{j=1}^{n} a_{ij}x_j$$

where $\vec{x}$ and $\vec{b}$ are column vectors

Notes

Row View of Matrix–Vector Product

Listing 1: Matrix–Vector Multiplication by Rows

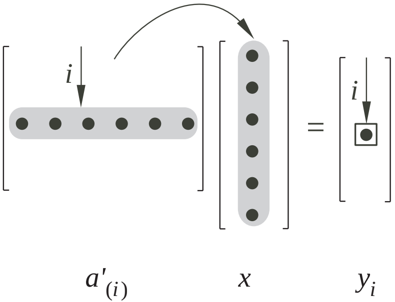
```
initialize:   b = zeros(m, 1)
for  i = 1, ..., m
   for  j = 1, ..., n
      b(i) = A(i, j)x(j) + b(i)
   end
end
```

Column View of Matrix–Vector Product

Listing 2: Matrix–Vector Multiplication by Columns

```
initialize:   b = zeros(m, 1)
for j = 1, . . . , n
   for i = 1, . . . , m
      b(i) = A(i, j)x(j) + b(i)
   end
end
```

Row View of Matrix–Vector Product



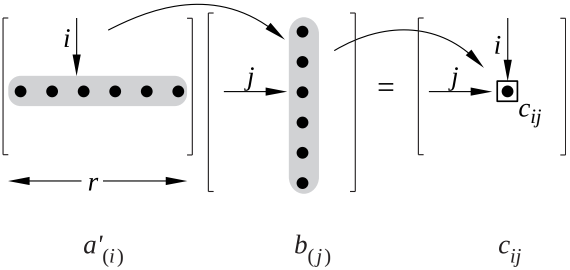
The row view generalizes very nicely to matrix–matrix multiplication.

Inner Product (Row) View of Matrix–Matrix Product

The product AB produces a matrix C. The c_{ij} element is the *inner product* of row i of A and column j of B.

$$AB = C \qquad \Longleftrightarrow \qquad c_{ij} = a'_{(i)} b_{(j)}$$

$a'_{(i)}$ is a row vector, $b_{(j)}$ is a column vector.



The inner product view of the matrix–matrix product is easier to use for hand calculations.

Matrix–Matrix Product Summary

The **Matrix–vector product** looks like:

$$\begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \begin{bmatrix} \bullet \\ \bullet \\ \bullet \end{bmatrix} = \begin{bmatrix} \bullet \\ \bullet \\ \bullet \\ \bullet \end{bmatrix}$$

The **vector–Matrix product** looks like:

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet \end{bmatrix} \begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} = \begin{bmatrix} \bullet & \bullet & \bullet \end{bmatrix}$$

Matrix–Matrix Product Summary

The **Matrix–Matrix product** looks like:

$$\begin{bmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{bmatrix} \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix} = \begin{bmatrix} \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet \end{bmatrix}$$

Matrix–Matrix Product Summary

Compatibility Requirement

$$\begin{array}{cccc} A & B & = & C \\ [m \times r] & [r \times n] & = & [m \times n] \end{array}$$

Inner dimensions must agree
Also, in general

$$AB \neq BA$$

Linear Systems as Matrix Equations

Linear System

$$\begin{cases} \phantom{x_1 - {}} 3x_2 + x_3 = 10 \\ x_1 - 2x_2 + x_3 = 6 \\ 3x_1 \phantom{{}- 2x_2} - x_3 = 13 \end{cases}$$

Matrix equation

$$\begin{bmatrix} 0 & 3 & 1 \\ 1 & -2 & 1 \\ 3 & 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 10 \\ 6 \\ 13 \end{bmatrix}$$

$$Ax = b$$

We know the "answer" (the vector $\vec{b}$).
Our goal is to work backwards and find $\vec{x}$.

Matrix Inverse

Let A be a square (i.e. $n \times n$) with real elements. The *inverse* of A is designated A^{-1}, and has the property that

$$A^{-1}A = I \qquad \text{and} \qquad A\,A^{-1} = I$$

The **formal solution** to $Ax = b$ is $x = A^{-1}b$.

$$\begin{aligned} Ax &= b \\ A^{-1}Ax &= A^{-1}b \\ Ix &= A^{-1}b \\ x &= A^{-1}b \end{aligned}$$

Matrix Inverse

- formal solution to $Ax = b$ is $x = A^{-1}b$

Matrix Inverse

- formal solution to $Ax = b$ is $x = A^{-1}b$
- BUT it is *bad* evaluate x this way

Matrix Inverse

- formal solution to $Ax = b$ is $x = A^{-1}b$
- BUT it is *bad* evaluate x this way
- why?

Matrix Inverse

- formal solution to $Ax = b$ is $x = A^{-1}b$
- BUT it is *bad* evaluate x this way
- why?
- we will not form A^{-1}, but solve for x directly using Gaussian elimination (or similar).

Formal Solution when A is $n \times n$

The *formal solution* to $Ax = b$ is

$$x = A^{-1}b$$

where A is $n \times n$.
If A^{-1} exists then A is said to be **nonsingular**.
If A^{-1} does not exist then A is said to be **singular**.

Formal Solution when A is $n \times n$

If A^{-1} exists then

$$Ax = b \qquad \Longrightarrow \qquad x = A^{-1}b$$

but

Do not compute the solution to $Ax = b$ by finding A^{-1}, and then multiplying b by A^{-1}!

We see: $x = A^{-1}b$

We do: **Solve $Ax = b$ by Gaussian elimination or an equivalent algorithm**

Why do we care as Numerical Analysts?

Open questions:

- How *expensive* is it to solve $Ax = b$?
- What problems (errors) will we encounter solving $Ax = b$?
- Some matrices are easy/cheap to use: diagonal, tridiagonal, etc.
  - are there others? what makes something a "good" matrix numerically?
  - are there bad ones? how do we identify them numerically?
- what do actual numerical analysts, engineers, developers, etc use?!?!

Singularity of A

If an $n \times n$ matrix, A, is **singular** then

- the columns of A are linearly dependent
- the rows of A are linearly dependent
- $\mathrm{rank}(A) < n$
- $\det(A) = 0$
- A^{-1} does not exist
- a solution to $Ax = b$ may not exist
- If a solution to $Ax = b$ exists, it is not unique

Know these conditions

Notes

Linear Independence

Two vectors lying along the same line are not independent

$$u = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} \qquad \text{and} \qquad v = -2u = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix}$$

Any two independent vectors, for example,

$$v = \begin{bmatrix} -2 \\ -2 \\ -2 \end{bmatrix} \qquad \text{and} \qquad w = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

define a plane. Any other vector in this plane of v and w can be represented by

$$x = \alpha v + \beta w$$

x is **linearly dependent** on v and w because it can be formed by a linear combination of v and w.

Linear Independence

A set of vectors is linearly independent if it is impossible to use a linear combination of vectors in the set to create another vector in the set. Linear independence is easy to see for vectors that are orthogonal, for example,

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

are linearly independent.

Linear Independence

Consider two linearly independent vectors, u and v.
If a third vector, w, *cannot* be expressed as a linear combination of u and v, then the set $\{u, v, w\}$ is linearly independent.

In other words, if $\{u, v, w\}$ is linearly independent then

$$\alpha u + \beta v = \delta w$$

can be true *only if* $\alpha = \beta = \delta = 0$.

More generally, if the only solution to

$$\alpha_1 v_{(1)} + \alpha_2 v_{(2)} + \cdots + \alpha_n v_{(n)} = 0 \tag{1}$$

is $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$, then the set $\{v_{(1)}, v_{(2)}, \ldots, v_{(n)}\}$ is **linearly independent**. Conversely, if equation (1) is satisfied by at least one nonzero α_i, then the set of vectors is **linearly dependent**.

Linear Independence

Let the set of vectors $\{v_{(1)}, v_{(2)}, \ldots, v_{(n)}\}$ be organized as the columns of a matrix. Then the condition of linear independence is

$$\left[\begin{array}{c|c|c|c} & & & \\ v_{(1)} & v_{(2)} & \cdots & v_{(n)} \\ & & & \end{array}\right] \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_n \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{2}$$

The columns of the $m \times n$ matrix, A, are linearly independent if and only if $x = (0, 0, \ldots, 0)^T$ is the only n element column vector that satisfies $Ax = 0$.

Summary of Requirements for Solution of $Ax = b$

Given the $n \times n$ matrix A and the $n \times 1$ vector, b

- the solution to $Ax = b$ exists and is unique for any b if and only if $\mathrm{rank}(A) = n$.

Recall: rank = # of linearly independent rows or columns

Recall: $Range(A)$ = set of vectors y such that $Ax = y$ for some x

Solving a system

$$Ax = b$$

Three situations:

1. A is nonsingular: There exists a unique solution $x = A^{-1}b$
2. A is singular and $b \in Range(A)$: There are infinite solutions.
3. A is singular and $b \notin Range(A)$: There no solutions.

1. $A = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 8 \end{bmatrix}$, then $x = \begin{bmatrix} 1/2 \\ 2 \end{bmatrix}$.
2. $A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, then infinitely many solutions. $x = \begin{bmatrix} 1/2 \\ \alpha \end{bmatrix}$.
3. $A = \begin{bmatrix} 2 & 0 \\ 0 & 0 \end{bmatrix}$ $b = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$, then no solutions.