1 Matrix-vector multiplication

$$\mathbf{b} = A\mathbf{x}$$

One may see this as showing $\mathbf{b}$ represented by a linear combination of the columns of A.

Let $\mathbf{x} \in \mathbb{C}^{\,n}$, A $\in \mathbb{C}^{\,m\times n}$
Then $\mathbf{b} \in \mathbb{C}^{\,m}$ (recalling that $\mathbf{b}$ = A$\mathbf{x}$), i.e.

$$\mathbf{b} = \underset{m\times 1}{\begin{bmatrix} b_1 \\ b_2 \\ \dots \\ b_m \end{bmatrix}} \qquad A\mathbf{x} = \underset{m\times n}{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}} \underset{n\times 1}{\begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}}$$

Specifically,

$$\underset{m\times 1}{\mathbf{b}} = \underset{m\times n}{A} \cdot \underset{n\times 1}{\mathbf{x}}$$

where

$$\mathbf{b} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \dots + x_n\mathbf{a}_n$$

where each $\mathbf{a}_i$ has length m (the number of rows in A).

Also: the map $\mathbf{x} \longmapsto$ A$\mathbf{x}$ is linear.

So $\mathbf{b} \in \mathbb{C}^{\,m}$ and

$$\mathbf{b}_i = \sum_{j=1}^{n} a_{ij}x_j \qquad i = 1, 2, \dots, m$$

Notice:
For any x, y $\in \mathbb{C}^{\,n}$ and $\alpha \in \mathbb{C}$,

$$A(\mathbf{x} + \mathbf{y}) = A\mathbf{x} + A\mathbf{y}$$

and

$$A(\alpha\mathbf{x}) = \alpha A\mathbf{x}$$

Conversely, any linear map $\mathbb{C}^n \longrightarrow \mathbb{C}^m$ can be expressed as multiplication by some m x n matrix.

1.1 Example: Vandermonde Matrix

Fix numbers x_1, x_2, ..., x_m.
If p and q are polynomials of degree < n and α is a scalar, then p + q and pα are polynomials of degree < n.

Note:

$$(p+q)x_i = p(x_i) + q(x_i)$$

and

$$(\alpha p)(x_i) = \alpha(p(x_i))$$

which implies that it is a linear transformation.

Let

$$A = \underset{m\times n}{\begin{bmatrix} 1 & x_1 & x_1^2 ... & x_1^{n-1} & \\ 1 & x_2 & x_1^2 ... & x_1^{n-1} & \\ 1 & x_3 & x_1^2 ... & x_1^{n-1} & \\ ... & ... & ... & ... & ... \\ 1 & x_m & x_m^2 ... & x_m^{n-1} & \end{bmatrix}}$$

If c is the column vector of coefficients of p, i.e.

$$c = \underset{n\times 1}{\begin{bmatrix} c_0 \\ c_1 \\ ... \\ c_{n-1} \end{bmatrix}}$$

and,

$$p(x) = c_0 + c_1 x + c_2 x^2 + ... + c_{n-1} x^{n-1}$$

then Ac gives the sampled polynomial values. **(what's the goal here?)**

1.2 Aside: vector spaces

A basis of $\mathbb{R}^n$ is composed of n vectors $\mathbf{v}_i \in \mathbb{R}^n$ such that:

$$\{\mathbf{v}_1, \mathbf{v}_2, ...\mathbf{v}_n\}$$

is a linearly independent set.

In practice, though, just because it's linearly independent, doesn't mean it's good to work with. If all the vectors are barely linearly independent, then it makes calculations hard.

Note: Chebychev and Lagandra polynomials are orthogonal polynomials (as opposed to the polynomials from the Vandermonde Matrix.)

1.3 Matrix Multiplication

Let

$$\underset{l\times n}{B} = \underset{l\times m}{A} \cdot \underset{m\times n}{C}$$

where each column of B is a linear combination of the columns of A. **(Is it also accurate to say each column of B is a linear combination of the rows of A?)** Specifically:

$$b_{ij} = \sum_{k=1}^{m} a_{ik}c_{kj}$$

Illustrated another way:

$$\underset{l\times n}{\begin{bmatrix} | & | & | & | \\ b_1 & b_2 & ... & b_n \\ | & | & | & | \end{bmatrix}} = \underset{l\times m}{\begin{bmatrix} | & | & | & | \\ a_1 & a_2 & ... & a_m \\ | & | & | & | \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} — & c_1 & — \\ — & c_2 & — \\ — & ... & — \\ — & c_n & — \end{bmatrix}}$$

For example,

$$\mathbf{b}_j = A\mathbf{c}_j = \sum_{k=1}^{m} c_{kj}a_k$$

1.4 Inner and Outer Products

Inner products (in the case of vectors **(is this true?)**) are dot products and they create scalars.
Outer products create matrices:

$$\underbrace{\begin{bmatrix} u_1 \\ u_2 \\ \dots \\ u_m \end{bmatrix}}_{m\times 1} \times \underbrace{\begin{bmatrix} v_1 & v_2 & \dots & n_n \end{bmatrix}}_{1\times n} = \underbrace{\begin{bmatrix} u_1v_1 & u_1v_2 & \dots & u_nv_1 \\ u_2v_1 & u_2v_2 & \dots & u_nv_2 \\ \dots & \dots & \dots & \dots \\ u_mv_1 & u_mv_2 & \dots & u_nv_m \end{bmatrix}}_{m\times n}$$

1.5 Range

The range of a matrix A, or range(A), is the set of all vectors that can be expressed as A**x** for some **x**.
You can think of this as a function that maps $\mathbb{C}^n$ to $\mathbb{C}^m$

Theorem:

range(A) is the space spanned by the columns of A.

Proof.

Recall:

$$Ax = b = \sum_{j=1}^{n} x_j a_j$$

and any Ax is a linear combination of columns of A.
This is saying that if something is in the range of A, then it is something spanned by the columns.
Conversely, any vector x in the space spanned by the columns of A can be written as a linear combination of the columns,

$$y = \sum_{j=1}^{n} x_j a_j$$

Forming a vector x with the coefficients x_j, we have

$$y = Ax$$

Thus, y ∈ range(A).

□

1.6 Null Space

The null space of A ∈ $\mathbb{C}^{m \times n}$, denoted null(A), is the set of vectors **x** such that

$$A\mathbf{x} = \mathbf{0} \in \mathbb{C}^m$$

We then write

$$\mathbf{0} = x_1 a_1 + x_2 a_2 + ... + x_n a_n$$

and **0** ∈ null(A)

1.7 Rank

The column rank of a matrix is the dimension of the space spanned by its columns.
The row rank of a matrix is the dimension of the space spanned by its rows.
In fact, the row rank of a matrix is always equal to its column rank (see: SVD).
Since these two values are equal, we just refer to the "rank" of the matrix.
A matrix is full rank if the rank is equal to min(m, n).

Theorem:

A matrix A ∈ $\mathbb{C}^{m \times n}$ with m ≥ n has full rank if and only if it maps no two distinct vectors to the same vector.

Proof.

$\Rightarrow$
If A has full rank (i.e. its rank is n), then its columns are linearly independent. Thus, its columns form a basis for range(A).
Thus, if $\mathbf{b} \in$ range(A), then $\mathbf{b}$ has a unique linear expansion in terms of the columns of A, and by

$$\mathbf{b} = A\mathbf{x} = \sum_{j=1}^{n} x_j a_j$$

every $\mathbf{b} \in$ range(A) has a unique $\mathbf{x}$ such that $\mathbf{b} = \mathrm{A}\mathbf{x}$.
$\Leftarrow$
Conversely, if A is not full rank, its columns a_j are dependent, and there is a nontrivial (not $\mathbf{0}$) combination such that

$$\sum_{j=1}^{n} c_j a_j = 0$$

Thus, A$\mathbf{c} = \mathbf{0}$ and then, for any $\mathbf{x}$,

$$A\mathbf{x} = A(\mathbf{x} + \mathbf{c})$$

$\square$

1.8 Inverse

A non-singular (or invertible) matrix is a square matrix with full rank.

The identity matrix I $\in \mathbb{C}^{\,n \times n}$:

$$I = \begin{bmatrix} | & | & | & | \\ e_1 & e_2 & ... & e_n \\ | & | & | & | \end{bmatrix}_{n \times n}$$

If AZ = I and ZA = I then Z = A^{-1}

Theorem:

For A $\in \mathbb{C}^{\,m \times m}$, the following are equivalent:

1. A has an inverse A^{-1}

2. rank(A) = m
3. rank(A) = $\mathbb{C}^{m}$
4. null(A) = $\{0\}$
5. 0 is not an eigenvalue of a.
6. 0 is not a singular value of a.
7. det A is not equal to 0

Warning: don't take the inverse and then find stuff. That's computationally expensive.
e.g. if $\mathbf{x} = \mathrm{A}^{-1}\mathbf{b}$ is the solution to $\mathrm{A}\mathbf{x} = \mathbf{b}$.

HW: problems 1.1 and 1.3