Linear algebra notes (based on Friedberg's text)

Abiyaz Chowdhury

April 3, 2018

1 Introduction

Definition 1.1 A vector space (or linear space) V over a field F consists of a set on which two operations (called addition and scalar multiplication, respectively) are defined so that for each pair of elements x, y in V there is a unique element $x + y$ in V, and for each element a in F, and each element $x \in V$ there is a unique element ax in V, such that the following conditions hold:

1. For all $x, y \in V$, $x + y = y + x$. (commutativity of addition)

2. For all $x, y, z \in V$, $(x + y) + z = x + (y + z)$. (associativity of addition)

3. There exists an element in V denoted by 0 such that $x + 0 = x$ for each $x \in V$. (identity under addition)

4. For each element $x \in V$ there is an element $y \in V$ such that $x + y = 0$. (existence of additive inverse)

5. For each element $x \in V$, $1x = x$. (identity under scalar multiplication)

6. For each pair of elements $a, b \in F$ and each element $x \in V$, $(ab)x = a(bx)$. (associativity under scalar multiplication)

7. For each element $a \in F$ and each pair of elements $x, y \in V$, $a(x + y) = ax + ay$. (scalar multiplication distributes over addition)

8. For each pair of elements $a, b \in F$ and each element $x \in V$, $(a + b)x = ax + bx$. (scalar multiplication distributes over addition)

The elements $x + y$ and ax are called the sum of x and y and the product of a and x, respectively. The elements of the field F are called scalars and the elements of the vector space V are called vectors. The element 0 as defined in the third item is called the zero vector of V.

Definition 1.2 An object of the form $(a_1, a_2, ..., a_n)$ where the entries $a_1, a_2, ..., a_n$ are elements of a field F, is called an n-tuple with entries from F. The elements $a_1, a_2, ..., a_n$ are called the entries or components of the n-tuple. Two n-tuples $(a_1, a_2...a_n)$ and $(b_1, b_2, .., b_n)$ with entries from a field F are called equal if $a_1 = b_i$ for all $i = 1, 2..., n$.

Definition 1.3 The set of all n-tuples with entries from a field F is denoted by F^n.

Definition 1.4 The sum of two n-tuples $(a_1, a_2...a_n)$ and $(b_1, b_2, .., b_n)$ in F^n yields the n-tuple $(c_1, c_2, .., c_n) \in F^n$ where $c_i = a_i + b_i$ for all $i = 1, 2, .., n$.

Definition 1.5 The scalar product of an n-tuple $(a_1, a_2...a_n)$ in F^n and a scalar $c \in F$ yields the n-tuple $(c_1, c_2, .., c_n) \in F^n$ where $c_i = c[a_i]$ for all $i = 1, 2, .., n$.

Theorem 1.1 F^n, under the operations defined above, is a vector space over F, with zero vector $(0, 0, ..0)$ where 0 is the additive identity of F.

Definition 1.6 An $m \times n$ matrix with entries from a field F is a rectangular array of the form

$$\begin{bmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ x_{d1} & x_{d2} & \ldots & x_{dn} \end{bmatrix}$$

where each entry $a_{ij}(1 \leq i \leq m, 1 \leq j \leq n)$ is an element of F. We call the entries a_{ij} with $i = j$ the diagonal entries of the matrix. The entries $a_{i1}, a_{2j}, ... a_{mj}$ compose the jth column of the matrix. The rows of the preceding matrix are regarded as vectors in F^n, and columns are regarded as vectors in F^m. The $m \times n$ matrix in which each entry equals zero is called the zero matrix and is denoted by O.

Definition 1.7 In the preceding matrix, m is the number of rows and n is the number of columns of the matrix. The entry of matrix A that lieds in row i and column j is denoted by A_{ij}. If a matrix has an equal number of rows and columns, it is called square. Two $m \times n$ matrices A and B are called equal if all their corresponding entries are equal, that is, if $A_{ij} = B_{ij}$ for $1 \leq i \leq m$ and $1 \leq j \leq n$.

Definition 1.8 The set of all $m \times n$ matrices with entries from a field F is denoted by $M_{m \times n}(F)$.

Definition 1.9 Addition of two $m \times n$ matrices A and B yields the $m \times n$ matrix C where $C_{ij} = A_{ij} + B_{ij}$.

Definition 1.10 Scalar Multiplication of an $m \times n$ matrix A with entries from F and a scalar $c \in F$ yields an $m \times n$ matrix C with entries from F where $C_{ij} = cA_{ij}$.

Theorem 1.2 $M_{m \times n}(F)$, under the operations defined above, is a vector space over F, with zero vector A where $A_{ij} = 0$ for all $i = 1, 2, .., m$ and $j = 1, 2, ..., n$ and 0 is the additive identity of F.

Definition 1.11 Matrix multiplication of an $m \times n$ matrix A and a $n \times p$ matrix B yields the $m \times p$ matrix C where $C_{ij} = \sum_{k=1}^{n} A_{ik} B_{kj}$.

Definition 1.12 Let S be any nonempty set and F be any field, and let $\mathcal{F}(S, F)$ denote the set of all functions from S to F. Two functions f and g in $\mathcal{F}(S, F)$ are called equal if $f(s) = g(s)$ for each $s \in S$. We define the sum of two functions $f, g \in \mathcal{F}(S, F)$ as the function $h \in \mathcal{F}(S, F)$ where $h(s) = f(s) + g(s)$. We define the scalar multiplication of a function $f \in \mathcal{F}(S, F)$ by a scalar $c \in F$ as the function $h \in \mathcal{F}(S, F)$ where $h(s) = c[f(s)]$.

Theorem 1.3 $\mathcal{F}(S, F)$ is a vector space, under the operations defined above, with zero vector $f \in \mathcal{F}(S, F)$ where $f(s) = 0$ for all $s \in 0$ is the zero vector in $\mathcal{F}(S, F)$.

Definition 1.13 A polynomial with coefficients from a field F is an expression of the form $f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0$ where n is a nonnegative integer and each a_k, called the coefficient of x^k, is in F. If $f(x) = 0$, that is, if $a_n = a_{n-1} = ... = a_0 = 0$, then $f(x)$ is called the zero polynomial, and the degree is defined as -1. Otherwise, the degree is defined as the largest exponent of x that appears in the representation $f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0$ with a nonzero coefficient. Two polynomials $f(x) = a_n x^n + a_{n-1} x^{n-1} + ... + a_1 x + a_0$ and $g(x) = b_n x^n + b_{n-1} x^{n-1} + ... + b_1 x + b_0$ are called equal if $m = n$ and $a_i = b_i$ for $i = 0, 1, .., n$.

Definition 1.14 A sequence in F is a function σ from the positive integers into F. A sequence is typically represented as an infinite series of numbers, i.e. $a_1, a_2, ...$ where the entries are elements in F. We can represent this sequence as simply $\{a_n\}$. The sum of two sequences $\{a_n\}$ and $\{b_n\}$ yields a sequence $\{c_n\}$ in F where $c_i = a_i + b_i$ and the scalar product of a sequence $\{a_n\}$ in F and a scalar $c \in F$ yields a sequence $\{c_n\}$ in F where $c_i = c[a_i]$.

Theorem 1.4 The set of all sequences in F, forms a vector space over F under the operations defined above.

2 Convex sets

3 Convex functions

Definition 3.1 A function $f : \mathbb{R}^n \to \mathbb{R}$ is convex if $\mathbf{dom}(f)$ is a convex set and if for all $x, y \in \mathbf{dom}(f)$ with $\theta \in [0, 1]$, we have $f(\theta x + (1 - \theta)y) \leq \theta f(x) + (1 - \theta)f(y)$.

Definition 3.2 A function $f : \mathbb{R}^n \to \mathbb{R}$ is strictly convex if $\mathbf{dom}(f)$ is a convex set and if for all $x, y \in \mathbf{dom}(f)$ with $x \neq y$ and $\theta \in (0, 1)$, we have $f(\theta x + (1 - \theta)y) < \theta f(x) + (1 - \theta)f(y)$.

Definition 3.3 f is concave if $-f$ is convex and f is strictly concave if $-f$ is strictly convex.

Theorem 3.1 No function can be both strictly concave strictly convex, but both convexity and concavity is possible. In fact, any function is affine if and only if it is both convex and concave.

Theorem 3.2 A function is convex if and only if it is convex when restricted to any line that intersects its domain.

Definition 3.4 A function $f : \mathbb{R}^n \to \mathbb{R}$ can be extended as $\widetilde{f} : \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by