Optimal Estimation University of Florida
Mechanical and Aerospace Engineering

HW 1

Issued: August 24, 2009, Due: August 31, 2009 (in class)

Problem 1 (Vector space). A vector space is a set that is closed under addition and scalar multiplication[1]. That is, V is a vector space if the following are true:

1. if x and y are elements of V, then $x + y \in V$.
2. if $x \in V$, then $ax \in V$, where a is a real number.

The elements of a vector space are called vectors. A familiar example of a vector space is $\mathbb{R}^n$, the n-dimensional real coordinate space. A subspace S of a vector space V is a subset of V that is a vector space in its own right. That is, $S \subseteq V$ and the elements of S satisfy the closure properties listed above. Consider the 2D plane (or, in fancy terminology, the vector space $\mathbb{R}^2$).

1. Show that the set of all 2-D vectors $x = [x_1, x_2]^T$ defined by $x_1 - x_2 = 0$ is a subspace, but that defined by $x_2 - x_1 = 1$ does not constitute a subspace.
2. Show that the set of real-valued continuous functions $f(x)$ that are defined on the interval $x \in [0, 1]$ form a vector space. (Function addition and scalar multiplication is defined as $(f + g)(x) = f(x) + g(x)$ and $(\alpha f)(x) = \alpha f(x)$.)

Problem 2 (A matrix as a linear mapping). An $m \times n$ real matrix A is best thought of as a linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$. For every n-dimensional vector x that you provide to the matrix as an "input", it produces the "output" $y \triangleq Ax$, which is an m-dimensional vector. We write $A : \mathbb{R}^n \to \mathbb{R}^m$. Compute Ax for

$$A = \begin{bmatrix} 1 & 2 \\ 5 & 4 \\ 13 & 6 \end{bmatrix},$$

and express the answer in terms of the two entries x_1, x_2 of the input vector x. Convince yourself that the output vector y is a linear combination of the columns of A, with the input x providing the coefficients (this is true for a general matrix, not just for this example, and is a very useful fact).

Problem 3 (Range and Null spaces). The range space of an $m \times n$ matrix A, denoted by $\mathcal{R}(A)$, is the set of all vectors in $\mathbb{R}^m$ that A can produce as output when you vary the input over all possible vectors $x \in \mathbb{R}^n$:

$$\mathcal{R}(A) = \{y | Ax = y \text{ for some } x\}.$$

The null space of a matrix A, denoted by $\mathcal{N}(A)$, is the set of all "input" vectors in $\mathbb{R}^n$ so that the output is 0:

$$\mathcal{N}(A) = \{x | Ax = 0\}.$$

---

[1] There are several additional technical details in the formal definition of a vector space, but we omit those here.

1. Show that $\mathcal{R}(A)$ is a vector space. What is the range space of $\begin{bmatrix}1 & 2\\ 2 & 4\\ 3 & 6\end{bmatrix}$? Show the range space pictorially. (Hint: use the "useful fact" from the previous problem)

2. Show that the null space is a vector space. What is the null space of the matrix $\begin{bmatrix}1 & 2\\ 2 & 4\\ 3 & 6\end{bmatrix}$? of $\begin{bmatrix}1 & 2.01\\ 2 & 4\\ 3 & 6\end{bmatrix}$? Show the null spaces pictorially.

Problem 4 (Span). The span of a set of vectors $L = \{x_1, \dots, x_\ell\}$ is the set of all vectors you can create by taking all possible linear combinations of them:

$$\mathit{span}(L) = \{y | y = \sum_{i=1}^{\ell} \alpha_i x_i\} \quad (\alpha_i\text{'s are scalars})$$

Denote the columns of an $m \times n$ matrix A by a_i, so that we can write $A = [a_1, a_2, \dots, a_n]$. Define the set $L = \{a_1, \dots, a_n\}$. What is another name for $\mathit{span}(L)$? (Hint: it has been defined earlier)

Problem 5 (Orthogonality). Two vectors x and y are said to be orthogonal if $x^T y = 0$. For a set of vectors S, its *orthogonal complement*, denoted by $S^\perp$ (read "S perp") is the set of vectors that are orthogonal to every vector in S:

$$S^\perp = \{y | x^T y = 0 \quad \forall x \in S\}$$

Two sets of vectors S_1 and S_2 are said to be orthogonal if every vector in S_1 is orthogonal to every vector in S_2.

1. Show that $S^\perp$ is a subspace (even if S is not).

2. In the familiar space $\mathbb{R}^3$, is the $x - y$ plane orthogonal to the $y - z$ plane?

3. *Orthogonal decomposition theorem*: Prove that for every matrix A,

$$\mathcal{R}(A)^\perp = \mathcal{N}(A^T) \quad \text{ and } \mathcal{N}(A)^\perp = \mathcal{R}(A^T)$$

   (Hint: to prove the first statement, start with the description:

$$\mathcal{R}(A)^\perp = \{x | x^T y = 0 \ \forall y \in \mathcal{R}(A)\}$$

   to prove the second statement from the first, use the fact that $S^{\perp^\perp} = S$ for a finite-dimensional vector space S.)

Problem 6 (Linear independence). A set of vectors $L = \{x_1, \dots, x_\ell\}$ is called linearly independent if the following is true:

$$x_1\alpha_1 + \dots + x_\ell\alpha_\ell = 0 \Rightarrow \alpha_1 = 0, \dots, \alpha_\ell = 0.$$

When we are dealing with vectors in $\mathbb{R}^n$, we can express this condition compactly. Since $x_i \in \mathbb{R}^n$, define a $n \times \ell$ matrix $X \triangleq [x_1, \dots, x_\ell]$ and a ℓ-vector $\alpha \triangleq [\alpha_1, \dots, \alpha_\ell]^T$. From the definition above, it follows that the columns of X are linearly independent if the only solution to $X\alpha = 0$ is the trivial solution $\alpha = 0$.

A linearly independent set of vectors that span a vector space V is called a *basis* of V. The unit vectors $S = \{e_1, e_2, \dots, e_n\}$ in $\mathbb{R}^n$ are called the standard basis of $\mathbb{R}^n$ (e_i has all zero entries except at the i^{th} location where it has a 1). Every vector space possesses a basis.

1. Find a basis for $\mathcal{R}(A)$, where $A = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 2 & 4 & 1 & 2 \end{bmatrix}$.
2. Convert this basis to an orthonormal basis, so that every vector element of the basis is orthogonal to every other vector element in it, and the norm of every vector element of the basis is 1. (Hint: use the following procedure: say x_1 and x_2 are the two vectors that form the basis. Decompose x_1 as $x_1 = x_1^o + \tilde{x}_1$ so that x_1^o is orthogonal to x_2. Now normalize x_1^o and call it $y_1 \triangleq \frac{x_1^o}{\|x_1^o\|}$. Notice that $\tilde{x}_1$ is linearly dependent with x_2, so it provides no information that x_2 does not provide and we can discard it. Normalize x_2 and call it y_2: $y_2 \triangleq \frac{x_2}{\|x_2\|}$. y_1 and y_2 span $\mathbb{R}^2$ (why?) and therefore form of a basis. Since they are orthogonal to each other and unit-norm, $\{y_1, y_2\}$ is an orthonormal basis of $\mathbb{R}^2$.)
3. How would you generalize this procedure of computing a orthonormal basis of $\mathbb{R}^n$ from a basis of $\mathbb{R}^n$? (It is called the Gram-Schmidt procedure)

Problem 7 (Positive definite matrices). A matrix P is said to be positive semi-definite if $x^T P x \geq 0$ for every $x \neq 0$. (P must be square for the definition to make sense, and $x = 0$ means every entry of x is 0). It is said to be positive definite if $x^T P x > 0$ for every $x \neq 0$. We write $P \geq 0$ to say P is positive semi definite, or $P > 0$ to say it is positive definite (**Caution!** $P > 0$ does not mean the elements of P are positive). Positive definite matrices are invertible. A symmetric positive definite matrix has a unique symmetric positive definite square root. (B is a square root of A if $BB = A$.)

Show that if H is an $m \times n$ matrix with full column rank (which is another way of saying that its columns are linearly independent), and P is a symmetric positive definite matrix, then $H^T P H$ is symmetric positive definite. Is this an if and only if statement?

Problem 8. A system of linear equations $Ax = y$ where A is a real $m \times n$ matrix, and y is a real m-vector are said to be *consistent* if it has a solution. It is inconsistent if there is no solution, that is, there is no x that satisfies $Ax = y$. If it is consistent, there are two possibilities, either there is a unique solution, or there are infinite number of solutions.

Show that the so-called system of *normal equations* $A^T A y = A^T x$ is consistent even if the original system $Ax = y$ is not. When is the solution to the normal equations unique?

(Hint: If S is a subspace of $\mathbb{R}^m$, every vector $y \in \mathbb{R}^m$ can be uniquely decomposed as $y = y_1 + y_2$, where $y_1 \in S$ and $y_2 \in S^\perp$.)

Problem 9. Define $\|x\|_2 \triangleq \sqrt{x_1^2 + \cdots + x_n^2} = x^T x$ for $x \in \mathbb{R}^n$. Prove the CBS (Cauchy-Bunyakovskii-Schwarz) Inequality for $\mathbb{R}^n$

$$|x^T y| \leq \|x\|_2 \|y\|_2.$$

(Hint: start with $\|\alpha x - y\|^2 \geq 0$ for a real α and choose α carefully.)

Problem 10 (**Norm**)**.** A norm in a vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ that maps elements of V into $\mathbb{R}$ that satisfies the following conditions:

1. $\|x\| \geq 0 \; \forall x \in V$ and $\|x\| = 0 \Leftrightarrow x = 0$.
2. $\|\alpha x\| = |\alpha| \|x\|$, α is a scalar.
3. $\|x + y\| \leq \|x\| + \|y\|$.

Show that the 2 norm of a vector in $\mathbb{R}^n$, that is, $\|x\|_2$ satisfies the above properties. Show the same for the 1-norm $\|x\|_1 \triangleq |x_1| + \cdots + |x_n|$.

Problem 11 (**Continuous-time to discrete-time transformation**)**.** Consider a plant whose model is given by the continuous time LTI system $\dot{x} = Ax + Bu, y = Cx$, where $x \in \mathbb{R}^n, u \in \mathbb{R}^m, y \in \mathbb{R}^p$. The input $u(t)$ to the plant is constructed by a ZOH DAC-board from the discrete samples u_k (so that $u(t) = u_k$ for $t \in [kT \; (k+1)T]$, where T is the sampling period) and the output is sampled by an ADC circuit to yield $y_k = y(kT)$. Derive the discrete-time version of this system $x_{k+1} = A_d x_k + B_d u_k, y_k = C_d x_k$, i.e., derive A_d, B_d, C_d in terms of A, B, C and T.

(hint: use the relationship $x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau$. Pay attention to the fact that you are dealing with vectors and matrices, not scalars.)