10.1 Linear Transformations and Norms.

A. Brief review of linear algebra.

1. Definition. A *linear transformation* $L: \mathbb{E}^n \to \mathbb{E}^m$ is a function with the property that for every $\mathbb{x},\ \mathbb{y} \in \mathbb{E}^n$, and scalars α, β,
$$L(\alpha\mathbb{x} + \beta\mathbb{y}) = \alpha L(\mathbb{x}) + \beta L(\mathbb{y}).$$
We denote the collection of all such linear transformations by $\mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$.

2. Definition. A *basis,* $\mathcal{B}$ in $\mathbb{E}^n$ is a collection of vectors $\mathcal{B} = \{\mathbb{b}_1, \mathbb{b}_2, \dots, \mathbb{b}_n\}$ that is linearly independent. If $\mathcal{B}$ is a basis then every vector $\mathbb{x} \in \mathbb{E}^n$ can be written uniquely as
$$\mathbb{x} = \sum_{j=1}^{n} \alpha_j\ \mathbb{b}_j$$
where the coefficients $\alpha_j = \mathbb{x} \cdot \mathbb{v}_j$ where $\{\mathbb{v}_j\}_{j=1}^{n}$ is the unique collection of vectors *biorthogonal* to $\mathcal{B}$, that is, such that $\mathbb{b}_j \cdot \mathbb{v}_k = 1$ if $j = k$ and 0 otherwise.
A basis $\mathcal{B}$ is *orthonormal* if $\mathbb{b}_j = \mathbb{v}_j$ for all j.

3. Remark. (a) $L: \mathbb{E}^1 \to \mathbb{E}^1$ is linear if and only if $L(x) = \alpha\, x$ for some real number α.

(b) $L: \mathbb{E}^1 \to \mathbb{E}^m$ is linear if and only if $L(x) = (\alpha_1 x, \alpha_2 x, \dots, \alpha_m x) = \mathbb{a}\, x$ for some fixed $\mathbb{a} \in \mathbb{E}^m$.

(c) $L: \mathbb{E}^n \to \mathbb{E}^1$ is linear if and only if $L(\mathbb{x}) = \mathbb{a} \cdot \mathbb{x} = \sum_{j=1}^{n} \alpha_j\, x_j$. In this case, $\alpha_j = L(\mathbb{e}_j)$ where $\mathbb{e}_j = (0, \dots, 0, 1, 0, \dots, 0)$ is the j^{th} element of the *standard basis* $\{\mathbb{e}_j\}_{j=1}^{n}$.

4. Theorem. Every linear transformation $L: \mathbb{E}^n \to \mathbb{E}^m$ can be represented as an $m \times n$ matrix.

B. The operator norm.

1. Definition. The *operator norm* (or just the *norm*) of $L \in \mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$ is defined by
$\|L\| = \inf\{C > 0: \|L\mathbb{x}\| \leq C\|\mathbb{x}\| \; for \; all \; \mathbb{x} \in \mathbb{E}^n\}$
or by

$$\|L\| = \left\{\sup \frac{\|L\mathbb{x}\|}{\|\mathbb{x}\|} : \mathbb{x} \in \mathbb{E}^n\right\}$$

2. Remark. (a) $\|L\mathbb{x}\| \leq C\|\mathbb{x}\|$ for all $\mathbb{x} \in \mathbb{E}^n$ says that the transformation L magnifies the norm of a given $\mathbb{x} \in \mathbb{E}^n$ by a factor of no more than C. The norm $\|L\|$ is the smallest such factor.

(b) The two quantities appearing in the definition of operator norm are the same. The proof of this is an exercise, but some of your work is done in Exercise 10.1.

(c) Given $L \in \mathcal{L}(\mathbb{E}^n, \mathbb{E}^m)$, $\|L\|$ is always finite, and in fact if L is represented by the matrix $[L] = (l_{i,j})_{m\times n}$ then

$$\|L\| \leq \left(\sum_{i=1}^{m}\sum_{j=1}^{n} l_{i,j}^2\right)^{1/2}$$

however, usually strict inequality holds.

(d) In fact, $||L||$ is independent of the matrix representation of L. Also $\det(L) = \det([L])$ is also independent of the matrix that is used to represent L.