IV.4. Diagonalizing Symmetric Operators

Note. We now consider eigenvalues, eigenvectors, and diagonalization of linear operators. In the main result (Theorem IV.4.5) we will see a result similar to one encountered in sophomore Linear Algebra.

Note. Recall that we can diagonalize an $n \times n$ matrix A if and only if $\mathbf{R}^n$ has a basis consisting of eigenvectors of A (see "Corollary 1. A Criterion for Diagonalization" in my online Linear Algebra notes on 5.2. Diagonalization). This is dealt with in connection to orthogonal matrices in:

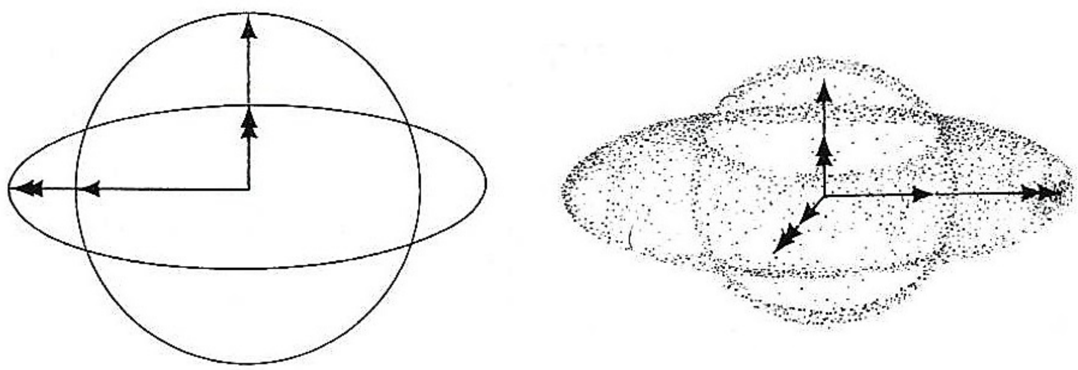
> **Fundamental Theorem of Real Symmetric Matrices.**
> Every real symmetric matrix A is diagonalizable. The diagonalization $D = C^{-1}AC$ can be achieved by using a real orthogonal matrix C.

See Theorem 6.8 of my online notes on 6.3. Orthogonal Matrices.

Note. We saw in the previous section that if metric vector space $(X, \mathbf{G})$ has an orthonormal basis then the matrix $[g_{ij}]$ is diagonal (and so is $[g^{ij}]$). Similarly if linear operator $\mathbf{A}$ has a collection of eigenvectors which are a basis β for X then it is easy to describe the behavior of $\mathbf{A}$. If $\mathbf{A}\mathbf{b}_i = \lambda_i \mathbf{b}_i$ for linearly independent $\mathbf{b}_i$ then

$$\mathbf{A}\mathbf{x} = \mathbf{A}(x^i \mathbf{b}_i) = x^i \mathbf{A}\mathbf{b}_i) = x^i \lambda_i \mathbf{b}_i,$$

or $\mathbf{A}(x^1, x^2, \ldots, x^n) = (\lambda_1 x^1, \lambda_2 x^2, \ldots, \lambda_n x^n)$ and $[\mathbf{A}]_\beta^\beta = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix}$. So geometrically we can break $\mathbf{A}$ down into scalar multiplication in various "directions." If the eigenvectors form an orthonormal basis (which is the case for a symmetric operator, as we'll see in Theorem IV.2.05), then we get geometric insight about $\mathbf{A}$ by considering the image of the unit sphere $\{\mathbf{x} \in X \mid \mathbf{x} \cdot \mathbf{x} = 1\}$ under $\mathbf{A}$. In some directions it is stretched (when $\lambda > 1$) and in some directions it is compressed (when $\lambda < 1$). This is illustrated in $\mathbb{R}^2$ and $\mathbb{R}^3$ in Figure 4.1 on page 93:



Definition IV.4.01. If $\mathbf{A}$ is a linear operator on an inner product space $(X, \mathbf{G})$, a vector $\mathbf{x}$ is *maximal* for $\mathbf{A}$ if $\mathbf{x}$ is a unit vector and $\mathbf{Ax} \cdot \mathbf{Ax} \geq \mathbf{Ay} \cdot \mathbf{Ay}$ for all unit vectors $\mathbf{y} \in X$.

Note. The inner product is a continuous mapping from X to $\mathbb{R}$ and in finite dimensions (our setting) the unit sphere is compact so a maximal vector exists.

Definition. If $\mathbf{x}$ is maximal for $\mathbf{A}$, then

$$\|\mathbf{Ax}\| = \max\{\mathbf{Ay}\cdot\mathbf{Ay} \mid \mathbf{y}\in X, \mathbf{y}\cdot\mathbf{y}=1\}$$

is the *norm* of $\mathbf{A}$, denoted $\|\mathbf{A}\|$.

Note. If we use the term "norm" then we should establish that $\|\cdot\|$ actually is a norm on the collection of linear operators on X. This is often called the "operator norm." See 2.4. Bounded Linear Operators in my online notes for Fundamentals of Functional Analysis (MATH 5740) for a few more details in a more general setting. In Exercise IV.4.1 it is to be shown that $\|\mathbf{Ay}\| \le \|\mathbf{A}\|\|\mathbf{y}\|$ (notice that this involves two different norms: the norm on X applies to $\mathbf{y}$ and $\mathbf{Ay}$, and the operator norm applied to $\mathbf{A}$).

Definition. A *symmetric* linear operator is a self-adjoint linear operator on an inner product space.

Note. We now consider eigenvalues and eigenvectors of symmetric operators.

Lemma IV.4.02. If $\mathbf{x}$ is a maximal vector of a symmetric operator $\mathbf{A}$ on an inner product space $(X, \mathbf{G})$ then $\mathbf{x}$ is an eigenvector of the operator $\mathbf{A}^2$, belonging to the eigenvalue $\|\mathbf{A}\|^2$.

Lemma IV.4.03. A symmetric operator $\mathbf{A}$ on a finite dimensional inner product space has an eigenvector belonging to an eigenvalue $+\|\mathbf{A}\|$ or $-\|\mathbf{A}\|$.

Lemma IV.4.04. If $\mathbf{X}$ is an eigenvector of a self-adjoint linear operator $\mathbf{A}$ on a metric vector space then $\mathbf{x} \cdot \mathbf{y} = 0$ implies $\mathbf{x} \cdot \mathbf{A}\mathbf{y} = 0$. That is, $\mathbf{A}(\mathbf{x}^\perp) \subseteq \mathbf{x}^\perp$ and so the map $\mathbf{y} \mapsto \mathbf{A}\mathbf{y}$ is an operator on $\mathbf{x}^\perp$, called the operator on $\mathbf{x}^\perp$ *induced* by $\mathbf{A}$.

Note. Now for out main result concerning symmetric linear operators. We use induction and perp spaces to give a fairly geometric argument.

Theorem IV.4.05. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space X, then X has an orthonormal basis of eigenvectors of $\mathbf{A}$.

Corollary IV.4.06. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space, then $[\mathbf{A}]_\beta^\beta$ is a diagonal matrix with respect to an orthonormal basis β. The diagonal entries of $[\mathbf{A}]_\beta^\beta$ are the (not necessarily distinct) eigenvalues of $\mathbf{A}$.

Note IV.4.A. Dodson and Poston deal with determinants of linear operators in Section I.03. However, they only do so by considering matrices. Of course, given a basis β we can produce a matrix $[\mathbf{A}] = [\mathbf{A}]_\beta^\beta$ representing linear operator $\mathbf{A}$. But different bases produce different matrices and different matrices will likely have different determinants. However, we can use a change of coordinates matrix to relate the matrices, say $[\mathbf{A}]_{\beta}^{\beta} = C_{\beta'}^{\beta}[\mathbf{A}]_{\beta'}^{\beta'}$. A change of coordinates matrix is invertible and so $\det(C_{\beta'}^{\beta}) \neq 0$. Hence

$$\det([\mathbf{A}]_{\beta}^{\beta}) = \det(C_{\beta'}^{\beta}[\mathbf{A}_{\beta'}^{\beta'}) = \det(C_{\beta'}^{\beta})\det([\mathbf{A}]_{\beta'}^{\beta'})$$

and we may have $\det([\mathbf{A}]_{\beta}^{\beta}) \neq \det([\mathbf{A}]_{\beta'}^{\beta'})$ (this can be accomplished by simply letting β' have the same vectors as β but with the first two vectors interchanged in which case the determinants differ by a factor of -1) but we still have $\det([\mathbf{A}]_{\beta}^{\beta}) = 0$ if and only if $\det([\mathbf{A}]_{\beta'}^{\beta'}) = 0$. So we can still find eigenvalues of linear operators on finite dimensional inner product spaces using the idea of a characteristic equation $\det([\mathbf{A} - \lambda\mathbf{I}]_{\beta}^{\beta}) = 0$ independent of the basis β! Dodson and Poston refer to the determinant of a linear operator (which, as we just argued, is not well defined); we will be more careful and indicate a basis and only take determinants of matrices with respect to a basis. As just observed, though, this is not necessary when looking for eigenvalues.

Corollary IV.4.07. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space with orthonormal basis β and if μ is a root of multiplicity m of the characteristic equation $\det([\mathbf{A} - \lambda\mathbf{I}]_{\beta}^{\beta}) = 0$ then the eigenspace belonging to μ has dimension m.

Note. In the inductive proof of Theorem IV.4.05, we repeatedly used Lemma IV.4.03 and hence at each step had real eigenvalues. So as a corollary to Theorem IV.4.05 we have the following.

Corollary IV.4.08. If $\mathbf{A}$ is a symmetric linear operator on a finite dimensional inner product space then all eigenvalues of $[\mathbf{A}]$ are real.

Note. The next result shows that there is an orthogonal basis with respect to a given symmetric bilinear form in an inner product space.

Corollary IV.4.09. In an inner product space $(X, \mathbf{G})$ for any symmetric bilinear form $\mathbf{h}$ on X we can find an orthonormal basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ for X such that $\mathbf{h}(\mathbf{b}_i, \mathbf{b}_j) = 0$ for $i \neq j$.

Definition IV.4.10. In inner product space $(X, \mathbf{G})$ and symmetric bilinear form $\mathbf{h}$ on X, an orthonormal basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ for X such that $\mathbf{h}(\mathbf{b}_i, \mathbf{b}_j) = 0$ for $i \neq j$ (which exists by Corollary IV.4.09) consists of the *principal directions* $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ of $\mathbf{h}$.

Note. The principal directions of symmetric bilinear form $\mathbf{h}$ may not by unique. For example, if μ is an eigenvalue of $\mathbf{A}_\mathbf{h}$ (see the proof of Theorem IV.4.09) of multiplicity 2 with corresponding unit eigenvector $\mathbf{b}_1$ and $\mathbf{b}_2$ (orthogonal), then $\mathbf{b}_1$ and $\mathbf{b}_2$ could be replaced with $(\mathbf{b}_1 + \mathbf{b}_2)/\sqrt{2}$ and $(\mathbf{b}_1 - \mathbf{b}_2)/\sqrt{2}$ which are also unit orthogonal vectors which span $\text{span}(\mathbf{b}_1, \mathbf{b}_2)$.

Definition. In inner product space $(X, \mathbf{G})$ with symmetric linear operator $\mathbf{h}$, if symmetric linear operator $\mathbf{A_h} : X \to X$ defined as $\mathbf{A_h}(\mathbf{x}) = \mathbf{G}_\uparrow(\mathbf{h_x})$ where $\mathbf{h_x} : X \to \mathbb{R}$ is defined as $\mathbf{h_x}(\mathbf{y}) = \mathbf{h}(\mathbf{x}, \mathbf{y})$ (see the proof of Theorem IV.4.09) has an eigenvalue λ such that $\mathbf{A_h}(\mathbf{x}) = \lambda \mathbf{x}$ for all $\mathbf{x} \in X$, then $\mathbf{h}$ is called *isotropic*.

Lemma IV.4.11. If $\mathbf{h}$ is isotropic, then $\mathbf{h} = \lambda \mathbf{G}$ for some $\lambda \in \mathbb{R}$ and $\mathbf{A_h} = \lambda \mathbf{I}$.

Note/Definition. So far, we have defined the space $\mathbb{H}^2$ and $\mathbb{H}^3$ with a Lorentz metric tensor in order to draw light cones in 2 and 3 dimensions. We refer to $\mathbb{R}^4$ with the metric tensor

$$(x^0, x^1, x^2, x^3) \cdot (y^0, y^1, y^2, y^3) = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3$$

as *Lorentz space* $\mathbb{L}^4$. Since the metric tensor is neither positive nor negative definite, then $\mathbb{L}^4$ is not an inner product space and so Theorem IV.4.05 does not hold (nor do its corollaries) for $\mathbb{L}^4$. In Exercise IV.4.5, a self-adjoint operator on $\mathbb{H}^2$ (not symmetric, since $\mathbb{H}^2$ is not an inner product space) is given which has only one eigenvalue but the corresponding eigenspace is dimension 1; so there is no basis of $\mathbb{H}^2$ consisting of eigenvectors of the operator. That is, Theorem IV.4.05 does not hold for metric vector space $\mathbb{H}^2$ (and similar problems exist for $\mathbb{H}^4$). The following result adds a hypothesis to self-adjoint linear operator $\mathbf{A}$ on $\mathbb{L}^4$ which implies the existence of an orthonormal basis of eigenvectors of $\mathbf{A}$.

Lemma IV.4.13. If a self-adjoint linear operator $\mathbf{A}$ on Lorentz space $\mathbb{L}^4$ has a timelike eigenvector $\mathbf{v}$ (i.e., $\mathbf{v} \cdot \mathbf{v} > 0$), then $\mathbb{L}^4$ has an orthonormal basis of eigenvectors of $\mathbf{A}$.

Revised: 5/24/2019