IV.3. Coordinates

Note. In this section we express $\mathbf{G}(\mathbf{x},\mathbf{y})$ in a metric vector space $(X,\mathbf{G})$ in terms of coordinates with respect to an ordered basis. This will give us a matrix of metric coefficients $[g_{ij}]$. We show that every metric vector space has an orthonormal basis (in Theorem IV.3.05). In "Sylvester's Law of Inertia" (Corollary IV.3.10) we see that, in a sense, every symmetric bilinear form on a metric vector space behaves similar to the Lorentz metric (in that it involves positive and negative coefficients).

Note IV.3.A. Let $(X,\mathbf{G})$ be a metric vector space. We denote, as usual, $\mathbf{G}(\mathbf{x},\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}$. Let $\beta = \{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n\}$ be an ordered basis for X. Suppose $\mathbf{x} = (x^1,x^2,\ldots,x^n)$ and $\mathbf{y} = (y^1,y^2,\ldots,y^n)$. Since $\mathbf{G}$ is bilinear, then with Einstein's summation convention we have:

$$\mathbf{G}(\mathbf{x},\mathbf{y}) = \mathbf{G}(x^i\mathbf{b}_i, y^j\mathbf{b}_j) = x^iy^j\mathbf{G}(\mathbf{b}_i,\mathbf{b}_j).$$

Note. Since $\mathbf{G}_\downarrow : X \to X^*$ is an isomorphism by Theorem IV.1.09, then it maps ordered basis $\beta = \{\mathbf{b}_1,\mathbf{b}_2,\ldots,\mathbf{b}_n\}$ of X to an ordered basis of X^*, $\mathbf{G}_\downarrow\beta = \{\mathbf{G}_\downarrow\mathbf{b}_1,\mathbf{G}_\downarrow\mathbf{b}_2,\ldots,\mathbf{G}_\downarrow\mathbf{b}_n\}$. Then the matrix that maps a coordinate vector $\mathbf{v}$ with respect to β to a coordinate vector $\mathbf{w}$ with respect to $\mathbf{G}_\downarrow\beta$ (where $\mathbf{w} = \mathbf{G}_\downarrow\mathbf{v}$, is $[\mathbf{G}_\downarrow]_\beta^{\mathbf{G}_\downarrow\beta} = [\delta_{ij}]$ (here we use double lower indices for Kronecker's delta function because the coordinate vectors will have components represented with upper indices).

Note IV.3.B. Let $\beta^* = \{\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^n\}$ be the dual basis (for X^*) of β. Then by definition (see Section III.1), $\mathbf{b}^j : X \to \mathbb{R}$ satisfies $\mathbf{b}^j(a^i\mathbf{b}_i) = a^j$. For $\mathbf{f} \in X^*$ where $\mathbf{f} - f_j\mathbf{b}^j = \mathbf{f}(\mathbf{b}_j)\mathbf{b}^j$. For $\mathbf{x} \in X$, we have $\mathbf{G}_\downarrow(\mathbf{x}) \in X^*$ so (with $\mathbf{f} = \mathbf{G}_\downarrow(\mathbf{x})$ we have $\mathbf{G}_\downarrow(\mathbf{x}) = \mathbf{G}_\downarrow(\mathbf{x})(\mathbf{b}_j)\mathbf{b}^j$. But $\mathbf{G}_\downarrow(\mathbf{x}) = \mathbf{x}^*$ where $\mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x}, \mathbf{y})$ by the definition of $\mathbf{G}_\downarrow$ (see Theorem IV.1.09), so $\mathbf{G}_\downarrow(\mathbf{x})(\mathbf{b}_j) = \mathbf{G}(\mathbf{x}, \mathbf{b}_j) = g_{ik}x^iy^k$ where $y^k = \begin{cases} 0 \text{ if } k \neq j \\ 1 \text{ if } k = j \end{cases}$ (by Note IV.3.A). That is, $\mathbf{G}_\downarrow(\mathbf{x})(_j) = g_{ij}x^i$ and this is the jth component of $\mathbf{G}_\downarrow(\mathbf{x})$ with respect to β^*. Hence $\mathbf{G}_\downarrow(\mathbf{x}) = g_{ij}x^i\mathbf{b}^j$. We now make a notational convention; we may denote vector $\mathbf{x} = (x^1, x^2, \ldots, x^n)$ as x^i. With this notation, we write $\mathbf{G}_\downarrow(\mathbf{x}) = g_{ij}x^i\mathbf{b}^j = g_{ij}x^i \in X^*$ (notice that summation is done over i so that index j determines the jth component of $\mathbf{G}_\downarrow(\mathbf{x})$ with respect to β^*). Also, $\mathbf{G}_\downarrow(\mathbf{x}) = g_{ij}x^j$. So with $\mathbf{x} = (x^1, x^2, \ldots, x^n)$ with respect to β then $\mathbf{G}_\downarrow(\mathbf{x}) = (g_{1j}x^j, g_{2j}x^j, \ldots, g_{nj}x^j)$ with respect to β^* and hence the matrix that converts $\mathbf{x}$ with respect to β to $\mathbf{G}_\downarrow(\mathbf{x})$ with respect to β^* is

$$[\mathbf{G}_\downarrow]_\beta^{\beta^*} = \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} = [g_{ij}].$$

Now $\mathbf{G}_\uparrow$ is the inverse of $\mathbf{G}_\downarrow$ so $[\mathbf{G}_\uparrow]_{\beta^*}^{\beta}$ is the inverse of $[\mathbf{G}_\downarrow]_\beta^{\beta^*}$. Let $[\mathbf{G}_\uparrow]_{\beta^*}^{\beta} = [g^{ij}]$. Then $g^{ij}g_{jk} = \delta_k^i$ and $g_{ki}g^{ij} = \delta_k^j$.

Note IV.3.C. If $\mathbf{x}, \mathbf{y} \in X^*$ where $\mathbf{x} = x_i$ and $\mathbf{y} = y_i$ with respect to β^* then

$$\begin{aligned} \mathbf{G}^*(\mathbf{x}, \mathbf{y}) &= \mathbf{G}(\mathbf{G}_\downarrow\mathbf{x}, \mathbf{G}_\uparrow\mathbf{y}) \text{ see Lemma IV.1.11} \\ &= \mathbf{G}(g^{ik}x_k\mathbf{b}_i, g^{j\ell}y_\ell\mathbf{b}_j) \text{ since } [g^{ij}] \text{ converts } \mathbf{x} \text{ and } \mathbf{y} \text{ coordinate vectors} \end{aligned}$$

$$\text{with respect to } \beta^* \text{ to } \mathbf{G}_\uparrow\mathbf{x} \text{ and } \mathbf{G}_\uparrow\mathbf{y} \text{ coordinate vectors}$$
$$\text{with respect to } \beta$$
$$= g^{ik}x_k g^{j\ell}y_\ell \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) \text{ since } \mathbf{G} \text{ is bilinear}$$
$$= g_{ij}g^{j\ell}x_k g^{j\ell}y_\ell \text{ since } g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j)$$
$$= (g_{ij}g^{j\ell})g^{ik}x_k y_\ell$$
$$= \delta_i^\ell g^{ik}x_k y_\ell \text{ since } [g_{ij}] \text{ and } [g^{j\ell}] \text{ are inverses}$$
$$= g^{\ell k}x_k y_\ell - g^{k\ell}x_k y_\ell \text{ since } g^{k\ell} = g^{\ell k}$$
$$= g^{ij}x_i y_j = x_i g^{ij} y_j.$$

In particular, for $\mathbf{x} = \mathbf{b}^i$ and $\mathbf{y} = \mathbf{b}^j$ (so $x^i = y^j = 1$ and all other coordinates of $\mathbf{x}$ and $\mathbf{y}$ with respect to β^* are 0) we have $\mathbf{G}^*(\mathbf{x}, \mathbf{y}) = \mathbf{G}^*(\mathbf{b}_i, \mathbf{b}_j) = g^{ij}$. That is, the components of $\mathbf{G}^*$ with respect to β^* are the g^{ij}. In terms of vectors and matrices, we can write

$$\mathbf{G}^*(\mathbf{x}, \mathbf{y}) = [x_1 \ x_2 \ \cdots \ x_n] \begin{bmatrix} g_{11} & g_{12} & \cdots & g_{1n} \\ g_{21} & g_{22} & \cdots & g_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ g_{n1} & g_{n2} & \cdots & g_{nn} \end{bmatrix} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}.$$

Note. We now shift out attention to finding a "nice" (i.e., orthonormal) basis for a matrix vector space. This should then yield nice matrices $[g^{ij}]$ and $[g_{ij}]$.

Definition IV.3.03. An *orthogonal set* in a metric vector space X is a subset S of X where for any $\mathbf{x}, \mathbf{y} \in S$ we have $\mathbf{x} \cdot \mathbf{x} \neq 0$, $\mathbf{y} \cdot \mathbf{y} = 0$, and $\mathbf{x} \cdot \mathbf{y} = 0$. An *orthonormal set* in X is an orthogonal set of unit vectors; i.e. $\mathbf{x} \cdot \mathbf{x} = 1$ for all $\mathbf{x} \in S$. An *orthonormal basis* for X is a basis which is an *orthonormal set.*

Lemma IV.3.04. For $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ an orthonormal basis for metric vector space $(X, \mathbf{G})$ in β coordinates we have $g_{ij} = \pm\delta_{ij}$.

Note. We know that every finite dimensional vector space has an orthonormal basis (since an n-dimensional vector space is isomorphic to $\mathbb{R}^n$ by the Fundamental Theorem of Finite Dimensional Vector Spaces and "orthonormal" is based on the usual dot product in $\mathbb{R}^n$). We wish to establish this result for metric vector space $(X, \mathbf{G})$ where "orthonormal" is based on metric tensor $\mathbf{G}$. The proof will use the Gram-Schmidt Process (though not by name). We first need a lemma.

Lemma IV.3.06. Nontrivial metric vector space $(X, \mathbf{G})$ possesses at least one non-null vector.

Theorem IV.3.05. Every metric vector space $(X, \mathbf{G})$ possess at least one orthonormal basis.

Note. By convention, we order an orthonormal basis $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n$ so that $\mathbf{b}_i \cdot \mathbf{b}_j = \begin{cases} +1 \text{ if } i \leq k \\ -1 \text{ if } i > k. \end{cases}$ With respect to this ordered basis,

$$\mathbf{x} \cdot \mathbf{y} = x^1 y^1 + x^2 y^2 + \cdots + x^k y^k - x^{k+1} y^{k+1} - x^{k+2} y^{k+2} - \cdots - x^n y^n.$$

The nest result show that parameter k does not depend on the choice of the basis.

Theorem IV.3.08. For any two orthonormal ordered bases $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ and $\beta^* = \{\mathbf{b}'_1, \mathbf{b}'_2, \ldots, \mathbf{b}'_n\}$ for a metric vector space $(X, \mathbf{G})$ with

$$\mathbf{b}_i \cdot \mathbf{b}_j = \begin{cases} +1 \text{ if } i \leq k \\ -1 \text{ if } i > k \end{cases} \text{ and } \mathbf{b}'_i \cdot \mathbf{b}'_j = \begin{cases} +1 \text{ if } i \leq \ell \\ -1 \text{ if } i > \ell, \end{cases}$$

we have $k = \ell$.

Note. We now see that in a metric vector space the number of "negative coefficients" and the number of "positive coefficients" in $\mathbf{G}$ (or in dot products or norms) is unique. This gives the following.

Corollary IV.3.09. Let $(X, \mathbf{G})$ be a matrix vector space with some basis $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$. Let $0 \leq k \leq n$ be orthonormal such that $\mathbf{b}_i \cdot \mathbf{b}_j = \begin{cases} +1 \text{ if } i \leq k \\ -1 \text{ if } i > k. \end{cases}$ Then the quantity $\sum_{i=1}^n g_{ii} = k(+1) + (n-k)(-1) = 2k - n$ is independent of the choice of the orthonormal basis.

Definition. The quantity $2k - n$ of Corollary IV.3.09 is the *signature* of $\mathbf{G}$.

Note. The following corollary shows that a result similar to Theorem IV.3.08 holds for any symmetric bilinear form.

Corollary IV.3.10. Sylvester's Law of Inertia.

Let $(X, \mathbf{G})$ be a metric vector space. For any symmetric bilinear form $\mathbf{F} : X \times X \to \mathbb{R}$, there is a choice of basis for which $\mathbf{F}$ has the form

$$\mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n)$$

$$= (x^1)^2 + (x^2)^2 + \cdots + (x^n)^2 - (x^{k+1})^2 - (x^{k+2})^2 - \cdots - (x^{k+\ell})^2$$

where $k + \ell \leq n$. Unless s or ℓ is zero, the subspace V^+ spanned by the basic vectors with $\mathbf{F}(\mathbf{b}_i, \mathbf{b}_j) = +1$ depends on the choice of basis; so does the subspace V^- spanned by those with $\mathbf{F}(\mathbf{b}_i, \mathbf{b}_i) = -1$. However, V^0, spanned by those with $\mathbf{F}(\mathbf{b}_i, \mathbf{b}_j) = 0$, depends only on $\mathbf{F}$, as do k and ℓ.

Lemma IV.3.11. Let $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ be a basis for $(X, \mathbf{G})$. Then the dual basis to β, $\beta^* = \{\mathbf{b}^1, \mathbf{b}^2, \ldots, \mathbf{b}^n\}$, is an orthonormal basis in the dual metric $\mathbf{G}^*$ on X^* if and only if β is an orthonormal basis for X.

Note. The proof of the following is left as Exercise IV.3.3.

Corollary IV.3.12. The signature of $\mathbf{G}^*$ equals the signature of $\mathbf{G}$.

Note. We now turn our attention to change of coordinate matrices (from one basis to another), first relating such matrices for operators A and A^T. Recall that we use "t" to represent the transpose of a matrix and "T" to represent the adjoint of a linear operator.

Lemma IV.3.13. If A is a linear operator on an inner product space $(X, \mathbf{G})$, then $[\mathbf{A}^T]_\beta^\beta = ([\mathbf{A}]_\beta^\beta)^t$ with respect to any orthonormal basis $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Note. Lemma IV.3.13 gives the matrix representation of the adjoint of $\mathbf{A}$ in terms of the matrix representation of $\mathbf{A}$ (and the two matrices are related by transpose). If $\mathbf{G}$ is indefinite, the relationship is more complicated.

Lemma IV.3.14. If $\mathbf{A}$ is a linear operator on a metric vector space $(X, \mathbf{G})$ then with respect to orthonormal basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ we have

$$[\mathbf{A}^T]_j^i = \left(\frac{g_{jj}}{g_{ii}}\right)[\mathbf{A}]_i^j = \left(\frac{g_{jj}}{g_{ii}}\right)[[\mathbf{A}]^t]_j^i$$

for $1 \leq i, j \leq n$, where $[\mathbf{A}]_j^i = a_{ij} = a_j^i$ is the entry in the ith row and jth column of $[\mathbf{A}]_\beta^\beta$ and there is no summation over i and j (though the Einstein convention implies it on the right hand side of the above equation).

Note. Since $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is an orthonormal basis in Lemma IV.2.14, then $g_{ii} = \mathbf{b}_i \cdot \mathbf{b}_i = \pm 1$ for each i. So the entries of $[\mathbf{A}^T]$ are the same is absolute value as the entries of $[\mathbf{A}]^t$. In fact, if $\mathbf{Q}$ is any linear operator on a metric vector space of dimension n and signature $\sigma = 2k - n$ then $\mathbf{Q}$ has a matrix representation with respect to β of the partitioned form

$$[\mathbf{Q}] = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

(where A is $k \times k$, B is $k \times (n-k)$, C is $(n-k) \times k$, and D is $(n-k) \times (n-k)$), where the first k columns are timelike (i.e., $g_{ii} = \mathbf{b}_i \cdot \mathbf{b}_i = +1$ for $1 \leq i \leq k$) and the last $n-k$ columns are spacelike (i.e., $g_{ii} = \mathbf{b}_i \cdot \mathbf{b}_i = -1$ for $k+1 \le i \le n$). So $\dfrac{g_{ii}}{g_{jj}} = +1$ for (1) $1 \le i \le k$ and $1 \le j \le n$, or (2) $k+1 \le i \le n$ and $k+1 \le j \le n$, and $\dfrac{g_{ii}}{g_{jj}} = -1$ for (1) $1 \le i \le k$ and $k+1 \le j \le n$ or (2) $k+1 \le i \le n$ and $1 \le j \le k$. So the matrix representation of $\mathbf{Q}^T$ has partitioned form

$$[\mathbf{Q}^T] = \left[\begin{array}{c|c} A^t & -C^t \\ \hline -B^t & D^t \end{array}\right].$$

Corollary IV.3.15. A linear operator $\mathbf{A}$ on a metric vector space is self-adjoint if and only if, with respect to an orthonormal basis, its matrix $[a_{ij} = [a^i_j]$ satisfies

$$a_{ij} = a^i_j = \frac{g_{jj}}{g_{ii}} a^j_i = \frac{g_{jj}}{g_{ii}} a_{ij}$$

for each i and j (there is not summation with respect to i or j here).

Note. If we apply Corollary IV.3.15 to an inner product space (where $\mathbf{G}$ is positive definite [or maybe negative definite]) then we see that $[\mathbf{A}]$ and $[\mathbf{A}^T]$ are just transposes of each other. So if $\mathbf{A}$ is self-adjoint then matrix $[\mathbf{A}]$ is symmetric. (This is based on the fact that we only consider real vector spaces. If we considered complex vector spaces then we would need to replace the transpose of a matrix with its conjugate transpose. See "Definition 9.2. Conjugate Transpose and Hermetian Adjoint" in my online Linear Algebra notes on 9.2. Matrices and Vector Spaces with Complex Scalars.)

Lemma IV.3.16. A linear operator **A** on an inner product space is orthogonal if and only if with respect to an orthonormal basis it has a matrix whose columns (respectively, rows) regarded as column (respectively, row) vectors form an orthonormal set in the standard inner product on $\mathbb{R}^n$.

Note. Since $\mathbb{R}^n$ is an inner product space with the usual dot product and every $n \times n$ matrix determines a linear operator on $\mathbb{R}^n$, then we can use Lemma IV.3.16 to establish the following result on matrices.

Corollary IV.3.17. The rows of a matrix are an orthonormal set (in the standard inner product) if and only if the columns are.

Revised: 5/14/2019