Differential Geometry

Chapter IV. Metric Vector Spaces

IV.3. Coordinates—Proofs of Theorems

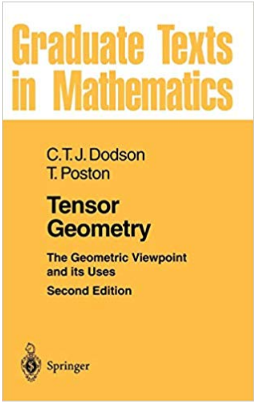

Table of contents

Lemma IV.3.04

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 $g_{ij} = \pm\delta_{ij}$.

Proof. By Note IV.3.A, we have $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j)$. Since β is an orthonormal set, for $i \neq j$ we have $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = 0$. In a metric vector space, $\|\mathbf{x}\| = \sqrt{|\mathbf{G}(\mathbf{x}, \mathbf{x})|}$, so we must have $|\mathbf{G}(\mathbf{b}_i, \mathbf{b}_j)| = 1$ or $\mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \pm 1$. Hence $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \pm\delta_{ij}$, as claimed. □

Lemma IV.3.04

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 $g_{ij} = \pm\delta_{ij}$.

Proof. By Note IV.3.A, we have $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j)$. Since β is an orthonormal set, for $i \neq j$ we have $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = 0$. In a metric vector space, $\|\mathbf{x}\| = \sqrt{|\mathbf{G}(\mathbf{x}, \mathbf{x})|}$, so we must have $|\mathbf{G}(\mathbf{b}_i, \mathbf{b}_j)| = 1$ or $\mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \pm 1$. Hence $g_{ij} = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \pm\delta_{ij}$, as claimed. □

Lemma IV.3.06

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

Proof. ASSUME not; i.e., assume $\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in X$. Then $(\mathbf{y} + \mathbf{z}) \cdot (\mathbf{y} + \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$ and so $\mathbf{y} \cdot \mathbf{y} + 2\mathbf{y} \cdot \mathbf{z} + \mathbf{z} \cdot \mathbf{z} = 0$ and $\mathbf{y} \cdot \mathbf{z} = -1/2(\mathbf{y} \cdot \mathbf{y} + \mathbf{z} \cdot \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$. But then $\mathbf{G}(\mathbf{y}, \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$ and $\mathbf{G}$ is not non-degenerate, a CONTRADICTION to the definition of metric vector space. So the assumption is false and hence there is some $\mathbf{x} \in X$ such that $\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = 0$. That is, there is some non-null $\mathbf{x} \in X$, as claimed. □

Lemma IV.3.06

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

Proof. ASSUME not; i.e., assume $\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = 0$ for all $\mathbf{x} \in X$. Then $(\mathbf{y} + \mathbf{z}) \cdot (\mathbf{y} + \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$ and so $\mathbf{y} \cdot \mathbf{y} + 2\mathbf{y} \cdot \mathbf{z} + \mathbf{z} \cdot \mathbf{z} = 0$ and $\mathbf{y} \cdot \mathbf{z} = -1/2(\mathbf{y} \cdot \mathbf{y} + \mathbf{z} \cdot \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$. But then $\mathbf{G}(\mathbf{y}, \mathbf{z}) = 0$ for all $\mathbf{y}, \mathbf{z} \in X$ and $\mathbf{G}$ is not non-degenerate, a CONTRADICTION to the definition of metric vector space. So the assumption is false and hence there is some $\mathbf{x} \in X$ such that $\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{x} \cdot \mathbf{x} = 0$. That is, there is some non-null $\mathbf{x} \in X$, as claimed. □

Theorem IV.3.05

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

Proof. By Lemma IV.3.06, there is non-null $\mathbf{x}_1 \in X$ such that $\mathbf{x}_1 \cdot \mathbf{x}_1 \neq 0$. Set $\mathbf{b}_1 = \dfrac{\mathbf{x}_1}{\|\mathbf{x}_1\|_{\mathbf{G}}}$. Now suppose that inductively for $1 \leq k < n$ we have $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ an orthonormal set in X. Let $B_k = \text{span}(\mathbf{B}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k)$. For $x \in B_k$, say $\mathbf{x} = x^i \mathbf{b}_i$ where $i = 1, 2, \ldots, k$, if $\mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{y} \in B_k$ then, in particular, $x^i = \mathbf{x} \cdot \mathbf{b}_i = 0$ for $i = 1, 2, \ldots, k$ and so $\mathbf{x} = \mathbf{0}$. That is, $\mathbf{G}$ is non-degenerate on B_k.

Theorem IV.3.05

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

Proof. By Lemma IV.3.06, there is non-null $\mathbf{x}_1 \in X$ such that $\mathbf{x}_1 \cdot \mathbf{x}_1 \neq 0$. Set $\mathbf{b}_1 = \dfrac{\mathbf{x}_1}{\|\mathbf{x}_1\|_{\mathbf{G}}}$. Now suppose that inductively for $1 \leq k < n$ we have $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ an orthonormal set in X. Let $B_k = \text{span}(\mathbf{B}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k)$. For $x \in B_k$, say $\mathbf{x} = x^i \mathbf{b}_i$ where $i = 1, 2, \ldots, k$, if $\mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{y} \in B_k$ then, in particular, $x^i = \mathbf{x} \cdot \mathbf{b}_i = 0$ for $i = 1, 2, \ldots, k$ and so $\mathbf{x} = \mathbf{0}$. That is, $\mathbf{G}$ is non-degenerate on B_k. By Corollary IV.2.05 $\dim(B_k^\perp) = n - k \neq 0$ and by Corollary IV.2.06 $\mathbf{G}$ is non-degenerate on $B_k^\perp$. So by Lemma IV.3.06 again, there is non-cull $\mathbf{x}_{k+1} \in B_k^\perp$. Set $\mathbf{b}_{k+1} = \dfrac{\mathbf{x}_{k+1}}{\|\mathbf{x}_{k+1}\|_{\mathbf{G}}}$. Then $\mathbf{b}_{k+1}$ is a unit vector orthogonal to $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$. That is, $\{\mathbf{b} - 1, \mathbf{b}_2, \ldots, \mathbf{b}_{k+1}\}$ is an orthonormal set. So by Mathematical Induction, there is an orthonormal set $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ in X.

Theorem IV.3.05

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

Proof. By Lemma IV.3.06, there is non-null $\mathbf{x}_1 \in X$ such that $\mathbf{x}_1 \cdot \mathbf{x}_1 \neq 0$. Set $\mathbf{b}_1 = \dfrac{\mathbf{x}_1}{\|\mathbf{x}_1\|_{\mathbf{G}}}$. Now suppose that inductively for $1 \leq k < n$ we have $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ an orthonormal set in X. Let $B_k = \text{span}(\mathbf{B}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k)$. For $x \in B_k$, say $\mathbf{x} = x^i \mathbf{b}_i$ where $i = 1, 2, \ldots, k$, if $\mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{y} \in B_k$ then, in particular, $x^i = \mathbf{x} \cdot \mathbf{b}_i = 0$ for $i = 1, 2, \ldots, k$ and so $\mathbf{x} = \mathbf{0}$. That is, $\mathbf{G}$ is non-degenerate on B_k. By Corollary IV.2.05 $\dim(B_k^\perp) = n - k \neq 0$ and by Corollary IV.2.06 $\mathbf{G}$ is non-degenerate on $B_k^\perp$. So by Lemma IV.3.06 again, there is non-cull $\mathbf{x}_{k+1} \in B_k^\perp$. Set $\mathbf{b}_{k+1} = \dfrac{\mathbf{x}_{k+1}}{\|\mathbf{x}_{k+1}\|_{\mathbf{G}}}$. Then $\mathbf{b}_{k+1}$ is a unit vector orthogonal to $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$. That is, $\{\mathbf{b} - 1, \mathbf{b}_2, \ldots, \mathbf{b}_{k+1}\}$ is an orthonormal set. So by Mathematical Induction, there is an orthonormal set $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ in X. By Exercise IV.3.2, $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is linearly independent and, since $\dim(X) = n$, is an orthonormal basis. □

Theorem IV.3.05

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

Proof. By Lemma IV.3.06, there is non-null $\mathbf{x}_1 \in X$ such that $\mathbf{x}_1 \cdot \mathbf{x}_1 \neq 0$. Set $\mathbf{b}_1 = \frac{\mathbf{x}_1}{\|\mathbf{x}_1\|_{\mathbf{G}}}$. Now suppose that inductively for $1 \leq k < n$ we have $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$ an orthonormal set in X. Let $B_k = \text{span}(\mathbf{B}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k)$. For $x \in B_k$, say $\mathbf{x} = x^i \mathbf{b}_i$ where $i = 1, 2, \ldots, k$, if $\mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} = 0$ for all $\mathbf{y} \in B_k$ then, in particular, $x^i = \mathbf{x} \cdot \mathbf{b}_i = 0$ for $i = 1, 2, \ldots, k$ and so $\mathbf{x} = \mathbf{0}$. That is, $\mathbf{G}$ is non-degenerate on B_k. By Corollary IV.2.05 $\dim(B_k^\perp) = n - k \neq 0$ and by Corollary IV.2.06 $\mathbf{G}$ is non-degenerate on $B_k^\perp$. So by Lemma IV.3.06 again, there is non-cull $\mathbf{x}_{k+1} \in B_k^\perp$. Set $\mathbf{b}_{k+1} = \frac{\mathbf{x}_{k+1}}{\|\mathbf{x}_{k+1}\|_{\mathbf{G}}}$. Then $\mathbf{b}_{k+1}$ is a unit vector orthogonal to $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_k$. That is, $\{\mathbf{b} - 1, \mathbf{b}_2, \ldots, \mathbf{b}_{k+1}\}$ is an orthonormal set. So by Mathematical Induction, there is an orthonormal set $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ in X. By Exercise IV.3.2, $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$ is linearly independent and, since $\dim(X) = n$, is an orthonormal basis. □

Theorem IV.3.08

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 $k = \ell$.

Proof. If $k = n$ then $\mathbf{G}$ is positive definite and so $\ell = n$. If $k = 0$ then $\mathbf{G}$ is negative definite and so $\ell = 0$. So, without loss of generality, we take $0 < k < n$. Let $N = \text{span}(\mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \ldots, \mathbf{b}_n)$. Then $\mathbf{G}$ is negative definite since for $\mathbf{x} = x^i \mathbf{b}_{k+i} \in N$ we have

$$\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{G}(x^i \mathbf{b}_{k+i}, x^i \mathbf{b}_{k+i}) = \sum_{i=1}^{n-k} -(x^i)^2.$$

Theorem IV.3.08

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 $k = \ell$.

Proof. If $k = n$ then $\mathbf{G}$ is positive definite and so $\ell = n$. If $k = 0$ then $\mathbf{G}$ is negative definite and so $\ell = 0$. So, without loss of generality, we take $0 < k < n$. Let $N = \text{span}(\mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \ldots, \mathbf{b}_n)$. Then $\mathbf{G}$ is negative definite since for $\mathbf{x} = x^i \mathbf{b}_{k+i} \in N$ we have

$$\mathbf{G}(\mathbf{x}, \mathbf{x}) = \mathbf{G}(x^i \mathbf{b}_{k+i}, x^i \mathbf{b}_{k+i}) = \sum_{i=1}^{n-k} -(x^i)^2.$$

Theorem IV.3.08 (continued 1)

Proof (continued). Let W be a subspace of X on which $\mathbf{G}$ is positive definite with basis $\omega = \{\omega_1, \omega_2, \ldots, \omega_r\}$. Consider the set $P = \{\omega_1, \omega_2, \ldots, \omega_r, \mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \ldots, \mathbf{b}_n\}$. If

$$a^1\omega_1 + a^2\omega_2 + \cdots + a^r\omega_r + a^{r+1}\mathbf{b}_{k+1} + a^{r+2}\mathbf{b}_{k+2} + \cdots + a^{r+(n-k)}\mathbf{b}_n = \mathbf{0}.$$

Then with the Einstein summation convention this implies

$$a^i\omega_i = -a^{r+j}\mathbf{b}_{k+j} \qquad (*)$$

and hence

$$(a^i\omega_i)\cdot(a^i\omega_i) = (-a^{r+j}\mathbf{b}_{k+j})\cdot(-a^{r+j}\mathbf{b}_{k+j}). \qquad (**)$$

But $\mathbf{G}$ (and so dot products) is nonnegative on W and nonpositive on N, so both sides of $(**)$ must be zero (since $a^i\omega_i \in W$ and $-a^{r+j}\mathbf{b}_{k+j} \in N$). Therefore both sides of $(*)$ are $\mathbf{0}$ (since the positive/negative definiteness of $\mathbf{G}$ implies from the dot product in $(**)$ that the constituent vectors in $(**)$ must be $\mathbf{0}$). So $a^i = 0$ for $i = 1, 2, \ldots, t + (n-k)$; that is, set P is a linearly independent set.

Theorem IV.3.08 (continued 1)

Proof (continued). Let W be a subspace of X on which $\mathbf{G}$ is positive definite with basis $\omega = \{\omega_1, \omega_2, \ldots, \omega_r\}$. Consider the set $P = \{\omega_1, \omega_2, \ldots, \omega_r, \mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \ldots, \mathbf{b}_n\}$. If

$$a^1\omega_1 + a^2\omega_2 + \cdots + a^r\omega_r + a^{r+1}\mathbf{b}_{k+1} + a^{r+2}\mathbf{b}_{k+2} + \cdots + a^{r+(n-k)}\mathbf{b}_n = \mathbf{0}.$$

Then with the Einstein summation convention this implies

$$a^i\omega_i = -a^{r+j}\mathbf{b}_{k+j} \qquad (*)$$

and hence

$$(a^i\omega_i) \cdot (a^i\omega_i) = (-a^{r+j}\mathbf{b}_{k+j}) \cdot (-a^{r+j}\mathbf{b}_{k+j}). \qquad (**)$$

But $\mathbf{G}$ (and so dot products) is nonnegative on W and nonpositive on N, so both sides of $(**)$ must be zero (since $a^i\omega_i \in W$ and $-a^{r+j}\mathbf{b}_{k+j} \in N$). Therefore both sides of $(*)$ are $\mathbf{0}$ (since the positive/negative definiteness of $\mathbf{G}$ implies from the dot product in $(**)$ that the constituent vectors in $(**)$ must be $\mathbf{0}$). So $a^i = 0$ for $i = 1, 2, \ldots, t + (n-k)$; that is, set P is a linearly independent set.

Theorem IV.3.08 (continued 2)

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 $k = \ell$.

Proof (continued). Since a linearly independent set of vectors in an n-dimensional vector space can have at most n elements, then $r + (n - k) \leq n$ or $r \leq k$ or $\dim(W) \leq k$ where W is an arbitrary subspace of X on which $\mathbf{G}$ is positive definite. Since $\mathbf{G}$ is positive definite on $\mathrm{span}(\mathbf{b}'_1, \mathbf{b}'_2, \ldots, \mathbf{b}'_\ell\}$ then we must have $\ell \leq k$. By a similar argument (interchanging the roles of β and β' and hence interchanging the roles of k and ℓ) we have $k \leq \ell$ and therefore $k = \ell$, as claimed. □

Corollary IV.3.10. Sylvester's Law of Inertia

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.

Proof. Set $V^N = \{\mathbf{x} \in X \mid \mathbf{F}(\mathbf{x}, \mathbf{y}) = 0 \text{ for all } \mathbf{y} \in X\}$. By linearity in the first term of $\mathbf{F}$, we see that V^N is closed under the vector sums and scalar multiplication and hence is a subspace of X. Let $\dim(V^N) = n - i$ and let $\{\mathbf{b}_{i+1}, \mathbf{b}_{i+2}, \ldots, \mathbf{b}_n\}$ be a basis of V^N. Extend this to a basis of X of the form $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n\}$ (which can be done by finding a basis for $(V^N)^\perp$ which is of dimension i by Corollary IV.2.05).

Corollary IV.3.10. Sylvester's Law of Inertia

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.

Proof. Set $V^N = \{\mathbf{x} \in X \mid \mathbf{F}(\mathbf{x}, \mathbf{y}) = 0 \text{ for all } \mathbf{y} \in X\}$. By linearity in the first term of $\mathbf{F}$, we see that V^N is closed under the vector sums and scalar multiplication and hence is a subspace of X. Let $\dim(V^N) = n - i$ and let $\{\mathbf{b}_{i+1}, \mathbf{b}_{i+2}, \ldots, \mathbf{b}_n\}$ be a basis of V^N. Extend this to a basis of X of the form $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n\}$ (which can be done by finding a basis for $(V^N)^\perp$ which is of dimension i by Corollary IV.2.05).

Corollary IV.3.10 (continued 1)

Proof (continued). Denote by W the subspace $\text{span}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i)$ of X. Let $\mathbf{w} \in W$. Then $\mathbf{F}(\mathbf{w}, \mathbf{v}) = 0$ for all $\mathbf{v} \in W$ implies

$$\mathbf{F}(\mathbf{w}, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots, x^i\mathbf{b}_i) + x^{i+1}\mathbf{F}(\mathbf{w}, \mathbf{b}_{i+1}) + x^{i+2}\mathbf{F}(\mathbf{w}, \mathbf{b}_{i+2}) + \cdots$$

$$+x^n\mathbf{F}(\mathbf{w}, \mathbf{b}_n) = 0$$

for all $(x^1, x^2, \ldots, x^n)$ since $x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i \in W$ and $\mathbf{b}_{i+1}, \mathbf{b}_{i+2}, \ldots, \mathbf{b}_n \in (V^N)^\perp$ and $W \subset (V^n)^\perp$. So by the bilinearity of $\mathbf{F}$ we have $\mathbf{F}(\mathbf{w}, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n) = 0$ for all $(x^1, x^2, \ldots, x^n)$ implies $\mathbf{F}(\mathbf{w}, \mathbf{x}) = 0$ for all $\mathbf{x} \in X$, which implies $\mathbf{w} \in V^N$ by the definition of V^N. But $\mathbf{w} \in W \subset (V^N)^\perp$. Since $\mathbf{w} \in V^N$ then $\mathbf{w} \in \text{span}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i)$. hence $\mathbf{w} = \mathbf{0}$. So the restricted symmetric bilinear form $\mathbf{F}|_{W \times W}$ is non-degenerate. So $(W, \mathbf{F})$ is a metric vector space and by Theorem IV.3.05 there is an orthonormal basis for W. Replace basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$ of W with this orthonormal basis of W (which we also denote as $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$).

Corollary IV.3.10 (continued 1)

Proof (continued). Denote by W the subspace span$(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i)$ of X. Let $\mathbf{w} \in W$. Then $\mathbf{F}(\mathbf{w}, \mathbf{v}) = 0$ for all $\mathbf{v} \in W$ implies

$$\mathbf{F}(\mathbf{w}, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots, x^i\mathbf{b}_i) + x^{i+1}\mathbf{F}(\mathbf{w}, \mathbf{b}_{i+1}) + x^{i+2}\mathbf{F}(\mathbf{w}, \mathbf{b}_{i+2}) + \cdots$$

$$+x^n\mathbf{F}(\mathbf{w}, \mathbf{b}_n) = 0$$

for all $(x^1, x^2, \ldots, x^n)$ since $x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i \in W$ and $\mathbf{b}_{i+1}, \mathbf{b}_{i+2}, \ldots, \mathbf{b}_n \in (V^N)^\perp$ and $W \subset (V^n)^\perp$. So by the bilinearity of $\mathbf{F}$ we have $\mathbf{F}(\mathbf{w}, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n) = 0$ for all $(x^1, x^2, \ldots, x^n)$ implies $\mathbf{F}(\mathbf{w}, \mathbf{x}) = 0$ for all $\mathbf{x} \in X$, which implies $\mathbf{w} \in V^N$ by the definition of V^N. But $\mathbf{w} \in W \subset (V^N)^\perp$. Since $\mathbf{w} \in V^N$ then $\mathbf{w} \in$ span$(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i)$. hence $\mathbf{w} = \mathbf{0}$. So the restricted symmetric bilinear form $\mathbf{F}|_{W \times W}$ is non-degenerate. So $(W, \mathbf{F})$ is a metric vector space and by Theorem IV.3.05 there is an orthonormal basis for W. Replace basis $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$ of W with this orthonormal basis of W (which we also denote as $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$).

Corollary IV.3.10 (continued 2)

Proof (continued). Then $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n\}$ is a basis of X for which $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$ is an orthonormal set and $\mathbf{F}(\mathbf{b}_p, \mathbf{b}_q) = 0$ whenever $1 \leq p \leq i$ and $i+1 \leq q \leq n$ by the definition of V^N. By Theorem IV.3.08, there is $0 \leq k \leq i$ such that $\mathbf{b}_j \cdot \mathbf{b}_j = \begin{cases} +1 \text{ if } j \leq k \\ -1 \text{ if } k < j \leq i. \end{cases}$ Then by the bilinearity of $\mathbf{F}$,

$$\begin{gathered}
\mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n) \\
= \mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_i) \\
+\mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^{i+1}\mathbf{b}_{i+1} + x^{i+2}\mathbf{b}_{i+2} + \cdots + x^n\mathbf{b}_n) \\
= \mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i) \\
+\mathbf{F}(x^{i+1}\mathbf{b}_{i+1} + x^{i+2}\mathbf{b}_{i+2} + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i) + 0 \\
= (x^1)^2 + (x^2)^2 + \cdots + (x^k)^2 - (x^{k+1})^2 - (x^{k+2})^2 - \cdots - (x^{k+\ell})^2
\end{gathered}$$

by the orthonormality of $\{x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i\}$, where $K + \ell = i$. So the basis exists, as claimed.

Corollary IV.3.10 (continued 2)

Proof (continued). Then $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i, \mathbf{b}_{i+1}, \ldots, \mathbf{b}_n\}$ is a basis of X for which $\{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_i\}$ is an orthonormal set and $\mathbf{F}(\mathbf{b}_p, \mathbf{b}_q) = 0$ whenever $1 \leq p \leq i$ and $i+1 \leq q \leq n$ by the definition of V^N. By Theorem IV.3.08, there is $0 \leq k \leq i$ such that $\mathbf{b}_j \cdot \mathbf{b}_j = \begin{cases} +1 \text{ if } j \leq k \\ -1 \text{ if } k < j \leq i. \end{cases}$ Then by the bilinearity of $\mathbf{F}$,

$$\mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n)$$
$$= \mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_i)$$
$$+\mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^n\mathbf{b}_n, x^{i+1}\mathbf{b}_{i+1} + x^{i+2}\mathbf{b}_{i+2} + \cdots + x^n\mathbf{b}_n)$$
$$= \mathbf{F}(x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i)$$
$$+\mathbf{F}(x^{i+1}\mathbf{b}_{i+1} + x^{i+2}\mathbf{b}_{i+2} + \cdots + x^n\mathbf{b}_n, x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i) + 0$$
$$= (x^1)^2 + (x^2)^2 + \cdots + (x^k)^2 - (x^{k+1})^2 - (x^{k+2})^2 - \cdots - (x^{k+\ell})^2$$

by the orthonormality of $\{x^1\mathbf{b}_1 + x^2\mathbf{b}_2 + \cdots + x^i\mathbf{b}_i\}$, where $K + \ell = i$. So the basis exists, as claimed.

Corollary IV.3.10 (continued 3)

Proof (continued). Now if k or ℓ is zero then $\mathbf{F}$ is, respectively, negative definite and positive definite, regardless of the basis used for $(W, \mathbf{F})$. When $k \neq 0 \neq \ell$, by Theorem IV.3.08, the choice of k (and hence the choice of ℓ) is independent of the basis of $(W, \mathbf{F})$. So for $k \neq 0 \neq \ell$, the dimensions of V^+ and V^- are determined by $\mathbf{F}$, but the spaces themselves depend on the choice of the basis since $V^+ = \text{span}(\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n)$ and $V^- = \text{span}(\mathbf{b}_{k+1}, \mathbf{b}_{k+2}, \ldots, \mathbf{b}_{k+\ell})$. Space $V^0 = V^N = \{\mathbf{x} \in X \mid \mathbf{F}(\mathbf{x}, \mathbf{y}) = 0 \text{ for all } \mathbf{y} \in X\}$ depends only on $\mathbf{F}$, as claimed. □

Lemma IV.3.11

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.

Proof. Basis β of X is orthonormal if and only if $\mathbf{G}(\mathbf{b}_i, \mathbf{b}_i) = \mathbf{b}_i \cdot \mathbf{b}_j = \pm\delta_{ij}$, which is equivalent to

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{bmatrix} = M$$

(provided we employ the convention of listing the -1's "first"). Now $[g_{ij}] = M$ if and only if the inverse $[g_{ij}]^{-1} = [g^{ij}] = M$. Then by Note IV.3.C, $\mathbf{G}^*(\mathbf{b}_i, \mathbf{b}_j) = g^{ij}$ so that $[g^{ij}] = M$ if and only if β^* is orthonormal. □

Lemma IV.3.11

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.

Proof. Basis β of X is orthonormal if and only if $\mathbf{G}(\mathbf{b}_i, \mathbf{b}_i) = \mathbf{b}_i \cdot \mathbf{b}_j = \pm\delta_{ij}$, which is equivalent to

$$[g_{ij}] = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -1 & 0 \\ 0 & 0 & \cdots & 0 & -1 \end{bmatrix} = M$$

(provided we employ the convention of listing the -1's "first"). Now $[g_{ij}] = M$ if and only if the inverse $[g_{ij}]^{-1} = [g^{ij}] = M$. Then by Note IV.3.C, $\mathbf{G}^*(\mathbf{b}_i, \mathbf{b}_j) = g^{ij}$ so that $[g^{ij}] = M$ if and only if β^* is orthonormal. □

Lemma IV.3.13

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 $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Proof. First, by definition, $\mathbf{G}_\downarrow \mathbf{b}_i = \mathbf{G}_\downarrow(\mathbf{b}_i) = \mathbf{x}^*$ where $\mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{y} \in X$. So

$$(\mathbf{G}_\downarrow \mathbf{b}_i)(\mathbf{b}_j) = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij}$$

for all $\mathbf{b}_j \in \beta$. Also for $\mathbf{b}^i \in \beta^*$ we have $\mathbf{b}^i(\mathbf{b}_j) = \mathbf{b}^i \mathbf{b}_j = \delta_{ij}$ for all $\mathbf{b}_j \in \beta$. So $\mathbf{G}_\downarrow(\mathbf{b}_i) = \mathbf{b}^i$ (since these elements of X^* are equal on basis β).

Lemma IV.3.13

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 $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Proof. First, by definition, $\mathbf{G}_\downarrow \mathbf{b}_i = \mathbf{G}_\downarrow(\mathbf{b}_i) = \mathbf{x}^*$ where $\mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x}, \mathbf{y})$ for all $\mathbf{y} \in X$. So

$$(\mathbf{G}_\downarrow \mathbf{b}_i)(\mathbf{b}_j) = \mathbf{G}(\mathbf{b}_i, \mathbf{b}_j) = \mathbf{b}_i \cdot \mathbf{b}_j = \delta_{ij}$$

for all $\mathbf{b}_j \in \beta$. Also for $\mathbf{b}^i \in \beta^*$ we have $\mathbf{b}^i(\mathbf{b}_j) = \mathbf{b}^i \mathbf{b}_j = \delta_{ij}$ for all $\mathbf{b}_j \in \beta$. So $\mathbf{G}_\downarrow(\mathbf{b}_i) = \mathbf{b}^i$ (since these elements of X^* are equal on basis β).

Lemma IV.3.13 (continued)

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 $\beta = \{\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_n\}$.

Proof (continued). Hence $[\mathbf{G}_\downarrow]_\beta^{\beta^*}$ is the identity matrix, and so is its inverse $[\mathbf{G}_\uparrow]_{\beta^*}^\beta$ (see Note IV.3.B). So

$$\begin{aligned}
[\mathbf{A}^T]_\beta^\beta &= [\mathbf{G}_\uparrow A^* \mathbf{G}_\downarrow]_\beta^\beta \text{ by definition of } \mathbf{A}^T \\
&= [\mathbf{G}_\uparrow]_{\beta^*}^\beta [\mathbf{A}^*]_{\beta^*}^{\beta^*} [\mathbf{G}_\downarrow]_\beta^{\beta^*} \\
&= [\mathbf{A}^*]_{\beta^*}^{\beta^*} \text{ since } [\mathbf{G}_\uparrow]_{\beta^*}^{\beta^*} = [\mathbf{G}_\downarrow]_\beta^{\beta^*} = \mathcal{I} \\
&= ([\mathbf{A}]_\beta^\beta)^t \text{ by Theorem III.1.A.}
\end{aligned}$$

□

Lemma IV.3.14

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

$$[\mathbf{A}^T]^i_j = \left(\frac{g_{jj}}{g_{ii}}\right)[\mathbf{A}]^j_i = \left(\frac{g_{jj}}{g_{ii}}\right)[[\mathbf{A}]^t]^i_j$$

for $1 \le i, j \le n$, where $[\mathbf{A}]^i_j = a_{ij} = a^i_j$ 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).

Proof. First $\mathbf{A}\mathbf{b}_j$ has coordinate vector with respect to β of

$$[a_{ij}][0 \;\; 0 \;\; \cdots \;\; 0 \;\; \underset{\substack{\uparrow \\ j}}{1} \;\; 0 \;\; \cdots \;\; 0 \;\; 0]^t = [a_{1j} \;\; a_{2j} \;\; \cdots \;\; a_{nj}]^t$$

so that $\mathbf{A}\mathbf{b}_j = a_{kj}\mathbf{b}_k$. So we have

$$(\mathbf{A}\mathbf{b}_j) \cdot \mathbf{b}_\ell = (a_{kj}\mathbf{b}_k) \cdot \mathbf{b}_\ell = a_{kj}(\mathbf{b}_k \cdot \mathbf{b}_\ell) = a_{\ell j}\mathbf{b}_\ell \cdot \mathbf{b}_\ell = [\mathbf{A}]^\ell_j(\mathbf{b}_\ell \cdot \mathbf{b}_\ell). \qquad (*)$$

Lemma IV.3.14

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

$$[\mathbf{A}^T]^i_j = \left(\frac{g_{jj}}{g_{ii}}\right)[\mathbf{A}]^j_i = \left(\frac{g_{jj}}{g_{ii}}\right)[[\mathbf{A}]^t]^i_j$$

for $1 \leq i, j \leq n$, where $[\mathbf{A}]^i_j = a_{ij} = a^i_j$ 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).

Proof. First $\mathbf{A}\mathbf{b}_j$ has coordinate vector with respect to β of

$$[a_{ij}][0\ \ 0\ \ \cdots\ \ 0\ \ \underset{\substack{\uparrow \\ j}}{1}\ \ 0\ \ \cdots\ \ 0\ \ 0]^t = [a_{1j}\ \ a_{2j}\ \ \cdots\ \ a_{nj}]^t$$

so that $\mathbf{A}\mathbf{b}_j = a_{kj}\mathbf{b}_k$. So we have

$$(\mathbf{A}\mathbf{b}_j)\cdot\mathbf{b}_\ell = (a_{kj}\mathbf{b}_k)\cdot\mathbf{b}_\ell = a_{kj}(\mathbf{b}_k\cdot\mathbf{b}_\ell) = a_{\ell j}\mathbf{b}_\ell\cdot\mathbf{b}_\ell = [\mathbf{A}]^\ell_j(\mathbf{b}_\ell\cdot\mathbf{b}_\ell). \qquad (*)$$

Lemma IV.3.14 (continued 1)

Proof (continued). Let $[\mathbf{A}^T]_\beta^\beta = [\mathbf{A}^T] = [c_{ij}]$. Then $\mathbf{A}^T\mathbf{b}_\ell = c_{k\ell}\mathbf{b}_\ell$. So

$$(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i = (c_{k\ell}\mathbf{b}_k)\cdot\mathbf{b}_i = c_{k\ell}(\mathbf{b}_k\cdot\mathbf{b}_i) = c_{i\ell}\mathbf{b}_i\cdot\mathbf{b}_i = [\mathbf{A}^T]_\ell^i(\mathbf{b}_i\cdot\mathbf{b}_i),$$

or $[\mathbf{A}^T]_\ell^i = \dfrac{(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i}$. Hence

$$\begin{aligned}
[\mathbf{A}^T]_\ell^i &= \frac{(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i} \\
&= \frac{(\mathbf{G}_\uparrow\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ by the definition of } \mathbf{A}^T \\
&= \frac{\mathbf{G}(\mathbf{G}_\uparrow\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell,\mathbf{b}_i)}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ since } \mathbf{G} \text{ determines dot products in } (X,\mathbf{G}) \\
&= \frac{(\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell)(\mathbf{b}_i)}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ since } \mathbf{G}_\uparrow(\mathbf{x}^*) = \mathbf{x} \text{ where } \mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x},\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}
\end{aligned}$$

for all $\mathbf{y}\in X$; see Theorem IV.1.09 and Note IV.1.A (here, $\mathbf{x}^* = \mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell$ and we take $\mathbf{y} = \mathbf{b}_i$)

Lemma IV.3.14 (continued 1)

Proof (continued). Let $[\mathbf{A}^T]_\beta^\beta = [\mathbf{A}^T] = [c_{ij}]$. Then $\mathbf{A}^T\mathbf{b}_\ell = c_{k\ell}\mathbf{b}_\ell$. So

$$(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i = (c_{k\ell}\mathbf{b}_k)\cdot\mathbf{b}_i = c_{k\ell}(\mathbf{b}_k\cdot\mathbf{b}_i) = c_{i\ell}\mathbf{b}_i\cdot\mathbf{b}_i = [\mathbf{A}^T]_\ell^i(\mathbf{b}_i\cdot\mathbf{b}_i),$$

or $[\mathbf{A}^T]_\ell^i = \dfrac{(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i}$. Hence

$$\begin{aligned}
[\mathbf{A}^T]_\ell^i &= \frac{(\mathbf{A}^T\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i} \\
&= \frac{(\mathbf{G}_\uparrow\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell)\cdot\mathbf{b}_i}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ by the definition of } \mathbf{A}^T \\
&= \frac{\mathbf{G}(\mathbf{G}_\uparrow\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell,\mathbf{b}_i)}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ since } \mathbf{G} \text{ determines dot products in } (X,\mathbf{G}) \\
&= \frac{(\mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell)(\mathbf{b}_i)}{\mathbf{b}_i\cdot\mathbf{b}_i} \text{ since } \mathbf{G}_\uparrow(\mathbf{x}^*) = \mathbf{x} \text{ where } \mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x},\mathbf{y}) = \mathbf{x}\cdot\mathbf{y}
\end{aligned}$$

for all $\mathbf{y}\in X$; see Theorem IV.1.09 and Note IV.1.A (here, $\mathbf{x}^* = \mathbf{A}^*\mathbf{G}_\downarrow\mathbf{b}_\ell$ and we take $\mathbf{y} = \mathbf{b}_i$)

Lemma IV.3.14 (continued 2)

Proof (continued). ...

$$
\begin{aligned}
[\mathbf{A}^T]^i_\ell &= \frac{(\mathbf{A}^*\mathbf{G}_\downarrow \mathbf{b}_\ell)(\mathbf{b}_i)}{\mathbf{b}_i \cdot \mathbf{b}_i} \\
&= \frac{(\mathbf{g}_\downarrow \mathbf{b}_\ell \mathbf{A})(\mathbf{b}_i)}{\mathbf{b}_i \cdot \mathbf{b}_i} \text{ since } \mathbf{A}^*(\mathbf{f}) = \mathbf{f} \circ \mathbf{A} \text{ by the definition of dual map} \\
&= \frac{(\mathbf{G}_\downarrow \mathbf{b}_\ell)(\mathbf{A}\mathbf{b}_\ell)}{\mathbf{b}_i \cdot \mathbf{b}_i} \text{ since function composition is associative} \\
&= \frac{\mathbf{b}_\ell \cdot (\mathbf{A}\mathbf{b}_i)}{\mathbf{b}_i \cdot \mathbf{b}_i} \text{ since } \mathbf{G}_\downarrow(\mathbf{x}) = \mathbf{x}^* \text{ where } \mathbf{x}^*(\mathbf{y}) = \mathbf{G}(\mathbf{x}, \mathbf{y}) = \mathbf{x} \cdot \mathbf{y} \\
&\qquad \text{for all } \mathbf{y} \in X\text{; see Theorem IV.1.09} \\
&= \frac{[\mathbf{A}]^\ell_i \mathbf{b}_\ell \cdot \mathbf{b}_\ell}{\mathbf{b}_i \cdot \mathbf{b}_i} \text{ by } (*) \\
&= [\mathbf{A}]^\ell_i \frac{g_{\ell\ell}}{g_{ii}},
\end{aligned}
$$

or $[\mathbf{A}^T]^i_j = [\mathbf{A}]^j_i \, g_{ii}/g_{jj}$, as claimed. □

Lemma IV.3.16

Lemma IV.3.16. A linear operator $\mathbf{A}$ 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 $\mathbb{R}^n$.

Proof. By Lemma IV.2.09, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. Let $[\mathbf{A}^T] = [c^i_j] = c_{ij}$. Then $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ is equivalent to $[\mathbf{A}^T] = [\mathbf{A}] = [\mathbf{I}]$ or $[\mathbf{A}^T]^k_i[\mathbf{A}^i_j = c_{ki}a_{ij} = c^k_i a^i_j = \delta^k_j$. By Lemma IV.3.13, $c_{ki} = c^j_i = a^i_k = a_{ik}$ and so the orthogonality of $\mathbf{A}$ is equivalent to $c_{ki}a_{ij} = a_{ik}a_{ij} = a^i_k a^i_j = \delta^k_j$. This is equivalent to the orthonormality of the columns of $[\mathbf{A}]$, as claimed.

Lemma IV.3.16

Lemma IV.3.16. A linear operator $\mathbf{A}$ 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 $\mathbb{R}^n$.

Proof. By Lemma IV.2.09, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. Let $[\mathbf{A}^T] = [c^i_j] = c_{ij}$. Then $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ is equivalent to $[\mathbf{A}^T] = [\mathbf{A}] = [\mathbf{I}]$ or $[\mathbf{A}^T]^k_i[\mathbf{A}^i_j = c_{ki}a_{ij} = c^k_i a^i_j = \delta^k_j$. By Lemma IV.3.13, $c_{ki} = c^j_i = a^i_k = a_{ik}$ and so the orthogonality of $\mathbf{A}$ is equivalent to $c_{ki}a_{ij} = a_{ik}a_{ij} = a^i_k a^i_j = \delta^k_j$. This is equivalent to the orthonormality of the columns of $[\mathbf{A}]$, as claimed.

By Corollary IV.2.10, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T$ is orthogonal. As argued above, $\mathbf{A}^T$ is orthogonal if and only if the columns of $[\mathbf{A}^T]$ are orthonormal. The columns of $[\mathbf{A}^T]$ are the rows of $[\mathbf{A}]$ by Lemma IV.3.13, so $\mathbf{A}$ is orthogonal if and only if the rows of $[\mathbf{A}]$ are orthonormal, as claimed. □

Lemma IV.3.16

Lemma IV.3.16. A linear operator $\mathbf{A}$ 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 $\mathbb{R}^n$.

Proof. By Lemma IV.2.09, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T\mathbf{A} = \mathbf{I}$. Let $[\mathbf{A}^T] = [c_j^i] = c_{ij}$. Then $\mathbf{A}^T\mathbf{A} = \mathbf{I}$ is equivalent to $[\mathbf{A}^T] = [\mathbf{A}] = [\mathbf{I}]$ or $[\mathbf{A}^T]_i^k[\mathbf{A}_j^i = c_{ki}a_{ij} = c_i^k a_j^i = \delta_j^k$. By Lemma IV.3.13, $c_{ki} = c_i^j = a_k^i = a_{ik}$ and so the orthogonality of $\mathbf{A}$ is equivalent to $c_{ki}a_{ij} = a_{ik}a_{ij} = a_k^i a_j^i = \delta_j^k$. This is equivalent to the orthonormality of the columns of $[\mathbf{A}]$, as claimed.

By Corollary IV.2.10, $\mathbf{A}$ is orthogonal if and only if $\mathbf{A}^T$ is orthogonal. As argued above, $\mathbf{A}^T$ is orthogonal if and only if the columns of $[\mathbf{A}^T]$ are orthonormal. The columns of $[\mathbf{A}^T]$ are the rows of $[\mathbf{A}]$ by Lemma IV.3.13, so $\mathbf{A}$ is orthogonal if and only if the rows of $[\mathbf{A}]$ are orthonormal, as claimed. □