Matrices and linear transformations

We work with a linear transformation $X \xrightarrow{A} Y$ and remind ourselves that for $\varphi(x) \in \mathbb{F}$ is $<\varphi \mid x>$.

Then $\mid Y >< \varphi$ stands for the function $X \rightarrow Y$ which acts on $x \in X$ by returning $\mid y >< \varphi \mid x >$ where we write

(i) $\mid x >, \mid y >$ etc. for vectors x, y in X and Y etc. calling them 'ket's and $<\varphi \mid, <\psi \mid$, etc. for forms φ, ψ in X' and Y' etc. calling them 'bra's so that $<\varphi \mid x>$ is a $bar(c)ket \in \mathbb{F}$ and $\mid x >< \varphi \mid$ is a ket-bra which is a linear transformation $X \rightarrow X$ returning $\mid x >< \varphi \mid w > \in X$ on input $\mid w > \in X$.

This is Dirac notion and terminology for linear algebra.

1.1 If $X \xrightarrow{|y><\varphi|} Y$ is not zero then

(i) $Im(\|y >< \varphi\|) =<< y >>$ *(* $\because$ *$v \in Im(\|y >< \varphi\|)$)* iff $v = \|y >< \varphi\|x >$ for some $x \in X$ i.e. a scalar multiple of $\|y >$; note that since we have here $< \varphi\| \neq 0$, we have $0 \neq \varphi_i =< \varphi\|e_i >$. for some basis vector e_i and then any $\lambda \in \mathbb{F}$ is $<\varphi \mid e_i>$ for the choice $x = \frac{e_i\lambda}{\varphi_i}$ since then $<\varphi \mid x >=< \varphi \mid \frac{e_i\lambda}{\varphi_i} >= \frac{<\varphi|e_i>}{\varphi_i}\lambda = \lambda$ so that $y\lambda =\mid y >< \varphi \mid \frac{e_i\lambda}{\varphi_i} >$ for each $y\lambda$ in $<< y >$.).

(ii) $ker\,(\mid y >< \varphi \mid) =<< \varphi >>^0$ *(recall $X'' = X$)* $(x \in ker\,(\mid y >< \varphi \mid)$ iff $\mid y >< \varphi \mid x >= 0$ iff $<\varphi \mid x >= 0$ since $\mid y > \neq 0$. This means, since $0 =< \varphi \mid x >=< x \mid \varphi >$, that $x \in << \varphi >>^0)$

(iii) Bilinearity: $\mid \lambda y + v >< \varphi \mid x >=\mid \lambda y >< \varphi \mid x > + \mid v >< \varphi \mid x >= [\lambda \mid y >< \varphi \mid + \mid v >< \varphi](x)$ at each $x \in X$ so that $\mid \lambda y + v >< \varphi \mid= \lambda \mid y >< \varphi \mid + \mid v >< \varphi \mid$, and $\mid >< \varphi + \varsigma\mu \mid x >=\mid y >< \varphi \mid x > + \mid y >< \varsigma \mid x > \mu = [\mid y >< \varphi \mid + \mu \mid y >< \varsigma \mid](x)$ at each x, so $\mid y >< \varphi + \varsigma\mu \mid =\mid y >< \varphi \mid + \mid >< \varsigma \mid \mu$ for each $y, v \in Y$, $\varphi, \varsigma \in X'$, $\lambda, \mu \in \mathbb{F}$

(iv) For $W \xrightarrow{A} X \xrightarrow{|><\varphi|} Y \xrightarrow{B} Z$ we have $B \circ (\mid y >< \varphi \mid x >=\mid By >< \varphi \mid x >$ at each $x \in X$ so that $B\circ \mid y >< \varphi \mid =\mid B(y) >< \varphi \mid (X \rightarrow X)$ and $(\because < \varphi \mid Aw >= ((\mid y >< \varphi \mid) \circ A)w$ at each $w \in X$ so that $(\mid y >< \varphi \mid) \circ A =\mid y >< A'\varphi' \mid (W \rightarrow Y)$

(v) Since $< A'\psi \mid x >=< \psi \mid Ax >$ for $X \xrightarrow{A} Y$, we have $< (\mid y >< \varphi \mid)'\psi \mid x >=< \psi \mid |y >< \varphi \mid x >$ at each $x \in X$ which means $(\mid y >< \varphi \mid)'(\psi) =< \psi \mid |y >< \varphi |>$

$=\mid \psi y >< \varphi \mid$

$= (\mid \varphi >< y \mid)(\psi) =\mid \varphi >< y \mid \psi >$

©Copyright Reserved IIT Delhi

(we have $X \xrightarrow{|y><\varphi|} X \xrightarrow{\psi} \mathbb{F}$, then by (iv), $\psi\circ \mid y >< \varphi \mid = \mid \psi(y) >< \varphi \mid$ but $\psi(y) =< \psi \mid y >=< y \mid \psi >$ since $y \in Y = Y''$ and thus $\psi(y)(\varphi) \in X'$ is simply $(< y \mid \psi >)\varphi$ which we are writing as $\mid \varphi >< y \mid \psi >$; note that $< y \mid \psi >=< \psi \mid y >\in \mathbb{F}$). at each $\psi\imath Y'$.

Thus we conclude:

$(|>< \varphi \mid)' = \mid \varphi >< y \mid$ *where* $\mid y >\in Y$ *on LHS has been written as* $< y \mid \in Y''$ *while* $< \varphi \mid \in X'$ *on LHS has been written as* $\mid \varphi > \in X'$ *as a ket since we have*

$Y' \xrightarrow{(|y><\varphi|)'=|><y|} X'$ *just as we have* $Z \xrightarrow{|w><\alpha|} W$ *for* $\mid w >\in Y$, $< \alpha \mid \in Z'$

(vi) For $X \xrightarrow{|y><\varphi|} Y \xrightarrow{|z><\psi|} Z$

$(\mid Z >< \psi \mid) \circ (\mid y >< \varphi \mid) =< \psi \mid y >\mid z >< \varphi \mid$ where $y \in Y$, $\psi \in Y'$, $\varphi \in X'$ and $x \in Z$ *(use* $B \circ (\mid y >< \varphi \mid) = \mid By >< \varphi \mid$ *established in (iv) with* $B = \mid z >< \psi \mid$*, noting that* $< \psi \mid y \in \mathbb{F}$*)*

(vii) For $X =<< e_1, \cdots, e_n >>$ so that $X' =<< e', \cdots, e' >>$ *(see ⊛ on page 10)*, we have $\sum_{i=1}^{n} \mid e_i >< e^i \mid e_j >= \mid e_j >$ which means $id_X = \sum_{i=1}^{n} \mid e_i >< e^i \mid$ (because $\sum_{i=1}^{n} \mid e_i >< e^i \mid$ is seem to be the function $e \to X$ given by $e_j \to e_j$) and $\sum_{i=1}^{n} \mid e_i >< e_i \mid e^j >= \mid e^j >$

$(\because < e_i \mid e^j >=< e^j \mid e_i >=$

$$\delta_{ij} == \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

)

so that $id'_X = \sum_{i=1}^{n} \mid e^i >< e_i \mid$.

(viii) Suppose $X =<< e_1, \cdots, e_n >>$, $Y =<< d_1, \cdots, d_m >>$. Given $X \xrightarrow{A} Y$, we have $A(e_i) \in Y$ and can write $A(e_i) = \sum_{j=1}^{m} \mid d_j > a_i^j$ for scalar $a_i^j \in \mathbb{F}$ uniquely; $a_i^j =< d^j \mid Ae_i >$. Then $A = A \circ id_X = A \circ (\sum_{i=1}^{n} \mid e_i >< e^i \mid$ *(see (vii) above)* $= \sum_{i=1}^{n} \mid Ae_i >< e^i \mid = \sum_{i=1}^{n}\sum_{j=1}^{m} \mid d_j > a_i^j < e^i \mid = \sum_{i=1}^{n}\sum_{j=1}^{m} a_i^j \mid d_j >< e^i \mid$ which means that the $m.n$ linear transformation $X \xrightarrow{|d_j><e^i|} Y$ which act on $x \in X$ to return the vector $\mid d_j >< e^i \mid x >= \mid d_j > X^i$ form a basis for the space $L(X, Y)$ since we just saw that any $X \xrightarrow{A} Y$ can be written uniquely as a linear combination $A = \sum_{i=1}^{n}\sum_{j=1}^{n} a_i^j \mid d_j >< e^i \mid$ in terms of these $\mid d_j >< e^i \mid$. The $m.n$ scalars a_j^j are written as an $m \times n$ matrix $a = [a_i^j]$ with i indexing the columns and j indexing the rows and this is called the matrix associated with the linear transformation $X \xrightarrow{A} Y$ with reference to the given indexed bases $e = \{e_i, \cdots, e_n\}$ of X and $d = \{d_1 \cdots, d_m\}$ of Y.

©Copyright Reserved IIT Delhi

Just how do we write this matrix? To begin, recall that when we say $y \in Y$ is uniquely written as $\sum_{j=1}^{m} d_j y^j$, the scalars y^j are $< d^j \mid y > \in \mathbb{F}$ where $\{d^1, \cdots, d^m\}$ is the basis of Y' associated with the basis $d = \{d_1, \cdots, d_m\}$ defined by $< d^j \mid d_k > :=$

$$\delta_{jk} = \begin{cases} 1 & \text{if } j = k \\ \\ 0 & \text{if } j \neq k \end{cases}$$

. Here $A(e_i) \in Y$ and so there scalars $(A(e_i))^j$ are $< d^j \mid Ae_i >$ which we have written as a_i^j; they depend on both $\{e_1, \cdots, e_n\}$ and $\{d_1, \cdots, d_m\}$ *(and of course on* $\{e^1, \cdots, e^n\}$ *and* $\{d^1, \cdots, d^n\}$*).* The matrix $a = [a_i^j]$ is obtained by writing the m-deep columns $A(e_i)$ as the i-the column and thus $a = \begin{pmatrix} a_1^1 & a_2^1 & \cdots & \cdots & a_n^1 \\ a_1^2 & a_2^2 & \cdots & \cdots & a_n^2 \\ \vdots & \vdots & \vdots & \vdots & \\ a_1^m & a_2^m & \cdots & \cdots & a_n^m \end{pmatrix}$ with a_i^j at the intersection of the j-th row and the i-the column.

(ix) The discussion in *(viii)* will be summarized by writing $(X, \underline{\underline{e}}) \xrightarrow[A]{a=[a_i^j]} (Y, \underline{\underline{\alpha}})$ and we keep writing $a_i^j =< d^j \mid Ae_i >$ so that we never forget that the matrix representation is in terms of the indexed bases $\underline{\underline{e}}$ and $\underline{\underline{d}}$ *(and of course their 'dual' or 'reciprocal' bases* $\{e^1, \cdots, e^n\}$, $\{d^1, \cdots, d^m\}$*)*; this is important because if the indexing is changed, the location of a_i^j is charged in the matrix and if either of the bases id changed, the entries will be different.

If $(X, \underline{\underline{e}} \xrightarrow[A]{a=[a_i^j]} (Y, \underline{\underline{d}}) \xrightarrow[B]{b=[b_j^k]} (Z, \underline{\underline{c}})$, $1 \leq i \leq n$, $1 \leq j \leq m$, $1 \leq k \leq p$, we have $(b \circ a)_i^k =< c^k \mid (B \circ A)e_i >=< C^k \mid B(\sum_{j=1}^{m} a_i^j \mid d_j >) >= \sum_{j=1}^{m} < C^k B d_j > a_i^j = \sum_{j=1}^{m} b_j^k a_i^j$ so that the composition of the linear transformation $B \circ A$ corresponds exactly to the composition of their matrix representation in the given bases.

Further, since $Y' =<< d^1, \cdots, d^m >>$ and $X' =<< e^1, \cdots, e^n >>$, we have the matrix representation of A' given by $Y' \xrightarrow[a']{A'} X'$ where $(a')_j^i =< e_i \mid A'(d^j) >=< A(e_i) \mid d^j >=< d^j \mid A(e_i) >= a_i^j$ *(note that we use* $Y'' = Y$*,* $X'' = X$*, and the fact that* $< \varphi \mid x >=< x \mid \varphi >$ *for kets* $\mid x >$ *and bras* $< \varphi \mid$*)*

©Copyright Reserved IIT Delhi

Also, $(X, \underline{\underline{e}}) \xrightarrow[I]{id_X} (X, \underline{\underline{e}})$ has $I_i^e = < e^l \mid id_X(e_i) > = < e^l \mid e_i >$

$$= \begin{cases} 1 & \text{if } l = i \\ 0 & \text{if } l \neq i \end{cases}$$

, $1 \leq i \leq n$, $1 \leq l \leq n$; thus the identity operator with respect to the same basis has the representation given by the identity*(square)* matrix which has on-diagonal entries 1 and off-diagonal entries 0.

We shall now supply some examples to illustrate the discussion. We use the standard terminology: Choice of a basis $\underline{\underline{e}}$ for X is called a coordinatization of X, in the representation $x = \sum e_i x^i = \sum \mid e_i > x^i$. the scalars x^i are called the coordinates or the components of the vector x with respect to the given coordinatization, and a linear transformation $X \xrightarrow{A} X$ will be frequently called a 'linear operator on X'.

2 Some illustrations.

2.1 Many authors do not wish to emphasize the need for indexing a basis, at least in the beginning. There is a reason for it. Consider the basis $\left\{ e_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, e_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \right\}$ for $\mathbb{F}^2$; the indexing would mean that $\{e_2, e_1\}$ is not a basis ($\{e_1, e_2\}$ but a different basis *(for this reason, ordered basis is usually the term for what we are calling indexed basis)*. To describe the action of a linear operator $\mathbb{F}^2 \xrightarrow{A} \mathbb{F}^2$ on basis vectors, one could say either (i) $Ae_1 = \alpha e_1$, $Ae_2 = \beta e_1 + \gamma e_2$ *(in which ordering the basis is irrelevant)*, or (ii) A has matrix $\begin{bmatrix} \alpha & \beta \\ 0 & \gamma \end{bmatrix}$ with respect to the ordered basis $\{e_1, e_2\}$ *(in which case ordering the basis is crucial, since one could easily but of course mistakenly, take this to mean $Ae_1 = \alpha e_1 + \beta e_2$, $Ae_2 = \gamma e_2$)*. The best thin to do seems to be: keep writing "$a_i^j = < d^j \mid Ae_i >, 1 \leq j \leq m\, 1 \leq i \leq n$ and $a = [a_i^j]$ is an $m \times n$ matrix for $(X, \underline{\underline{e}}) \xrightarrow{A} (Y, \underline{\underline{d}})$: $Ae_i = \sum \mid d_j > a_i^j$". Linear algebra is better understood without matrices but the subject is so computational in application that de-emphasizing matrices is almost scandalous.

2.2 (i) Find the matrix of $\mathbb{F}^3 \xrightarrow{A} \mathbb{F}^2$ given by $A \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 2x + 3y - z \\ 4x - y + 2z \end{pmatrix}$ relative to the basis

©Copyright Reserved IIT Delhi

$$\left\{\underbrace{\begin{pmatrix}1\\1\\0\end{pmatrix}}_{=b_1},\underbrace{\begin{pmatrix}1\\2\\3\end{pmatrix}}_{=b_2},\underbrace{\begin{pmatrix}1\\3\\5\end{pmatrix}}_{=b_3}\right\} \text{ of } \mathbb{F}^3 \text{ and } \left\{\underbrace{\begin{pmatrix}1\\2\end{pmatrix}}_{=c_1},\underbrace{\begin{pmatrix}2\\3\end{pmatrix}}_{=c_2}\right\} \text{ of } \mathbb{F}^2.$$

Hint:

That the two sets given are indeed bases can be verified by checking them for linear independence. For instance, we know that $\mathbb{F}^3$ *has three elements in any basis and the system* $\lambda_1 b_1 + \lambda_2 b_2 + \lambda_3 = 0$ *has no nontrivial solutions(verify) so that* $\{b_1, b_2, b_3\}$ *are linearly independent and must form a basis of* $\mathbb{F}^3$*. Next, if* $\begin{pmatrix}a\\b\end{pmatrix} \in \mathbb{R}^2$ *, solving* $x\begin{pmatrix}1\\2\end{pmatrix} + y\begin{pmatrix}2\\3\end{pmatrix} = \begin{pmatrix}a\\b\end{pmatrix}$ *supplies* $x = -3a + 2b$, $y = 2a - b$ *so that in the given basis* $\{c_1, c_2\}$ *of* $\mathbb{F}^2$*, we must have this vector as* $\begin{pmatrix}-3a+2b\\2a=b\end{pmatrix}$ *and directly calculating therefore, we have* $Ab_1 = \begin{pmatrix}5\\3\end{pmatrix} = -9c_1 + 7c_2$ $Ab_2 = c_1 + 2c_2$*,* $Ab_3 = 4c_1 + c_2$*. The desired matrix will be obtained by these column vectors* Ab_1*,* Ab_2*,* Ab_3 *so that it is* $a = \begin{pmatrix}-9 & 1 & 4\\7 & 2 & 1\end{pmatrix}$*. The vector* $\begin{pmatrix}2\\5\\-3\end{pmatrix}$ *of* $\mathbb{F}^3$ *is calculated to be* $\begin{pmatrix}11\\-21\\12\end{pmatrix}$ *in the basis* $\{b_1, b_2, b_3\}$ *(verify that any* $\begin{pmatrix}a\\b\\c\end{pmatrix} \in \mathbb{F}^3$ *in the basis* $\{b_1, b_2, b_3\}$ *will be* $\begin{pmatrix}-a+2b-c\\5a-5b+2c\\-3a+3b-3c\end{pmatrix}$*) and calculating* $A\begin{pmatrix}2\\5\\-3\end{pmatrix} = \begin{pmatrix}4+15+3\\8-5-6\end{pmatrix} = \begin{pmatrix}22\\-3\end{pmatrix}$ *by the supplied formula, writing this as* $\begin{pmatrix}-66-6\\44+3\end{pmatrix}$ *is the basis* $\{c_1, c_2\}$*, we get* $\begin{pmatrix}-72\\47\end{pmatrix}$ *which is precisely* $\begin{pmatrix}-9 & 1 & 4\\7 & 2 & 1\end{pmatrix}\begin{pmatrix}11\\-21\\12\end{pmatrix}$ *as it should be and displays the action of* A *as matrix action.*

(ii) For $\mathbb{F}^3 \xrightarrow{A} \mathbb{F}^2$ given by $A\begin{pmatrix}x\\y\\z\end{pmatrix} = \begin{pmatrix}3x+2y-4z\\x-5y+3z\end{pmatrix}$ with

$B_1 = \left\{ \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{=u_1}, \underbrace{\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}}_{=u_2}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{=u_3} \right\}$ and $B_2 = \left\{ \underbrace{\begin{pmatrix} 1 \\ 3 \end{pmatrix}}_{=v_1}, \underbrace{\begin{pmatrix} 2 \\ 5 \end{pmatrix}}_{=v_2} \right\}$ as bases, the representation of A is $\begin{pmatrix} -7 & -33 & -13 \\ 4 & 19 & 8 \end{pmatrix}$. Verify this and the action displayed on $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ to be $\begin{pmatrix} -13x - 20y + 26z \\ 8x + 11y - 15z \end{pmatrix}$ in the coordinatizon B_2 for $A\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ as given.

2.3 Prove that if $X \xrightarrow{A} Y$ is a linear transformation, there exists a basis $\underline{\underline{e}}$ for X and a basis $\underline{\underline{d}}$ for Y for which A has the matrix representation. $a = \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ where I_r is the identity matrix of order r and r is the rank of A; take $dim\, X = n < \infty$, $dim\, Y = m < \infty$.

Hint:

we know that $r = dim\,(A(\alpha))$ hence $dim(ker\,A) = n - r$ ((ii) on page 8). Let $\{e_{r+1}, \cdots, e_n\}$ be a basis for $ker\,A$ then as a linearly independent subset of X, this can be extended to a basis $\underline{\underline{e}} = \{e_1, \cdots, e_r, e_{r+1}, \cdots, e_n\}$ of X. Write $d_j = A(e_j)$ for $1 \le j \le r$. Since $X \xrightarrow{A} A(X)$ is surjective, it must be an isomorphism $X \xrightarrow{A} A(X)$ ((d) on page 9) and thus these r vectors $d_j \in A(X)$ will form a basis for the r-dimensional space $A(X)$; then they are linearly independent in Y and we can extend it to a basis $\underline{\underline{d}} = \{d_1, \cdots, d_m\}$ of Y (read the HINT for (ii) on page 8 if you are uncertain about the process of getting $\underline{\underline{e}}$ and $\underline{\underline{d}}$). Then

$d_1 = A(e_1) = 0.d_1 + 0.d_2 + \cdots + 0.d_m$

$d_2 = A(e_2) = 0.d_1 + 1.d_2 + \cdots + 0.d_m$

$\cdots \quad \cdots \quad \cdots$

$d_r = A(e_r) = 0.d_1 + \cdots + 1.d_r + 0.d_{r+1} \cdots + 0.d_m$

$0 = A(e_{r+1}) = 0.d_1 + \cdots + 0.d_m \; (\because e_{r+1} \in ker\, A)$

$\cdots \quad \cdots \quad \cdots$

$0 = a(en) = 0.d_1 + \cdots + 0.d_m \; (\because e_n \in ker\, A)$

©Copyright Reserved IIT Delhi

and therefore $(X, \underline{\underline{e}}) \xrightarrow[a]{A} (Y, \underline{\underline{d}})$ *must be a* $= \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix}$ *as advertized.*

2.4 What is $dim\,(L(X,Y))$ if $dim\,X = n < \infty$, $dim\,Y = m < \infty$?

Hint:

Fix a basis $\underline{\underline{e}} = \{e_1, \cdots, e_n\}$ *for* X *and a basis* $\underline{\underline{d}} = \{d_1, \cdots, d_m\}$ *for* Y. *Then the corresponding* $(X, \underline{\underline{e}}) \xrightarrow[a]{A} (Y, \underline{\underline{d}})$, $A \leftrightarrow a$, *is bijective. Since the vector space of all* $m \times n$*matrices(over* $\mathbb{F}$*) is of dimension* $m.n$ *(with basis* $\begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \end{pmatrix}$, $\begin{pmatrix} 0 & 1 & 0 & \cdots \\ 0 & . & . & \cdots \end{pmatrix}$, $\begin{pmatrix} 0 & \cdots & 0 \\ 1 & \cdots & . \\ 0 & . & \cdots \end{pmatrix}$ *etc; prove this),* $dim\,(L(X,Y)) = m.n$.

$\|$ *But it is far better to simply appeal to the fact that* $\{| \, d_j >< e_i \, | \, |1 \leq i \leq n \, 1 \leq j \leq m\}$ *forms a basis for* $L(X,Y)$ *(see page 15)*

2.5 If $X \times X \xrightarrow{B} \mathbb{F}$ is such that $B(\lambda x + u, w) = \lambda B(x.w) + B(u,w)$ and $B(x, u + w\lambda) = B(x,u) + B(x,w)\lambda$, we say it is a bilinear form on X *(compare 1.3 page 2)*. Show that if B and C are bilinear forms on X, $(B+C)(x,u) := B(x,u) + C(x,u)$ makes the collection of all bilinear forms on X a vector space*(with the obvious scalar multiplication)*. Call this $\underline{\underline{\text{Bilim}(X)}}$.

(i) If $\underline{\underline{e'}} = \{e'_1, \cdots, e'_n\}$ is a basis of X' show that $\beta_i^j(x,u) := \beta'_j(x)\beta'_i(u)$ forms a basis of $\underline{\underline{Bilim\,(X)}}$.

Hint:

If $B \in \underline{\underline{Bilim\,(X)}}$ *and* $b_i^j := B(e_j, e_i) \in \mathbb{F}$ *where* $\underline{\underline{e}} = \{e_1, \cdots, e_n\}$ *is the corresponding basis for* $X = X''$, *we get* $(\sum\limits_{i,j} b_i^j \beta_i^j)(e_s, e_t) = \sum\limits_{i,j} b_i^j \beta_i^j(e_s, e_t) = \sum b_i^i e'_j(e_s) e'_i(e_t) = \sum b_i^j \delta_{js}\delta_{it} = b_t^s = B(e_s, e_t)$; *since* $\{(e_s, e_t) \mid 1 \leq s \leq n, \, 1 \leq t \leq n\}$ *forms a basis of* $X \times X$, *we get all of* B *from this; thus the* β_i^j *span* $\underline{\underline{Bilim\,X}}$*(verify of course that* β_i^j *are bilinear). If* $B = \sum b_i^j \beta_i^j = 0$, *we have* $B(e_s, e_t) = b_s^t = 0$ *for each* s, t *which means this set is linearly independent and is thus a basis for* $\underline{\underline{Bilin\,(X)}}$. *In particular dim* $(\underline{\underline{Bilin\,(X)}} = (dim\,X)^2$.

(ii) Entering $b_i^j = B(e_j, e_j)$ into n^2 matrix at the intersection of the j-th row and i-th column will get us a matrix b which is called the matrix of the bilinear form B with respect to the basis $\underline{\underline{e}}$ of X. For example, the bilinear form on $\mathbb{F}^2$ given by $B\left(\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}\right) := 2x_1y_1 - 3x_1y_2 + x_2y_2$,

©Copyright Reserved IIT Delhi

with respect to the basis $\left\{ e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ has the matrix $\begin{pmatrix} 2 & -1 \\ 2 & 0 \end{pmatrix}$ *(Verify this).*

(iii) Given a Bilinear form $X \times X \xrightarrow{B} \mathbb{F}$ on X, the function $X \xrightarrow{Q} \mathbb{F}$ supplied by $Q(x) := B(x, x)$ is called the quadratic form associated to B.

Show that the correspondence established by *(ii)* above, $B \leftrightarrow b$ is bijective in the sense that $B(x, u) = x^t bu$ *(x^t is the transpose of the column vector x)* and that any n^2-matrix b will raise a bilinear form $B(x, u) := x^t bu$. Further, show that $Q \leftrightarrow q$ establishes a similar correspondence between quadratic forms Q and symmetric n^2-matrices q via $Q(x, x) := x^t qx$.

(iv) Show that if B raises the quadratic form Q, we can get the bilinear form B from

$2B(u, v) = B(u, u) + B(u, v) + B(v, u) + B(v, v) - B(u, u) - B(v, v)$

$= B(u + v, u + v) - B(u, u) - f(v, v)$

$= Q(u + v) - Q(u) - Q(v)$

that is, $B(u, v)$ can be defined from the quadratic form Q via $B(u, v) := \frac{1}{2}[Q(u+v) - Q(u) - Q(v)]$ (provided of course that $2 \neq 0$ in $\mathbb{F}$ i.e. $\mathbb{F} \neq \mathbb{Z}_2$)

Note:

Frequently, to save space, $\begin{pmatrix} x^1 \\ x^2 \end{pmatrix} \in \mathbb{F}^n$ will be written $[x^1 . \cdots x^n]^t$ or $(x^1, \cdots, x^n)$.

I am faithfully following linear algebra and Group Representation *(Volume I)* by Ronald Shaw *(Academic Press 1982).* This handout covers the selection from the first *(25)* pages from the first chapter. There is a copy in our central library but it is perhaps easier to work through the handout compared to the some what terse presentation of the book.

©Copyright Reserved IIT Delhi