MATH2040A Week 6 Tutorial Notes

1 Change of Coordinates

1.1 Review

Recall that on two ordered bases $\beta = \{ v_1, \dots, v_n \}, \gamma = \{ w_1, \dots, w_m \}$ of finite dimensional vectors spaces V, W respectively (with $n = \dim(V)$ and $m = \dim(W)$), we have the representation maps $[\cdot]_\beta$ on vectors and $[\cdot]_\beta^\gamma$ on linear maps:

- for $v \in V$, $[v]_\beta \in F^n$ is a column vector that is formed by the linear combination $v = \sum([v]_\beta)_i v_i$. Similar for γ and W
- for $T \in L(V, W)$, $[T]_\beta^\gamma \in F^{m\times n}$ is a matrix that is formed by horizontally stacking the column vectors $[T(v_j)]_\gamma$ together, or equivalently by the relation $T(v_j) = \sum_i([T]_\beta^\gamma)_{ij} w_i$

and they satisfy

- $[Tv]_\gamma = [T]_\beta^\gamma [v]_\beta$
- $[TU]_\alpha^\gamma = [T]_\beta^\gamma [U]_\alpha^\beta$
- T is invertible if and only if $[T]_\beta^\gamma$ is invertible, in which case $[T^{-1}]_\gamma^\beta = \left([T]_\beta^\gamma\right)^{-1}$

The representations under different choices of bases are related by the *change of coordinate* matrices:

Definition 1.1. On two ordered bases β_1, β_2 of the same vector space, $[\mathrm{Id}]_{\beta_1}^{\beta_2}$ is the *change of coordinate matrix* from β_1-coordinate to β_2-coordinate.

Using the properties we have

$$\begin{aligned} [v]_{\beta_2} &= [\mathrm{Id}]_{\beta_1}^{\beta_2}[v]_{\beta_1} \\ [T]_{\beta_2}^{\gamma_2} &= [\mathrm{Id}_W]_{\gamma_1}^{\gamma_2}[T]_{\beta_1}^{\gamma_1}\left([\mathrm{Id}_V]_{\beta_1}^{\beta_2}\right)^{-1} \\ [\mathrm{Id}]_{\beta_2}^{\beta_1} &= \left([\mathrm{Id}]_{\beta_1}^{\beta_2}\right)^{-1} \end{aligned}$$

If we want to find the change of coordinate matrix $[\mathrm{Id}]_\beta^\gamma$ for given ordered bases β, γ, we can

- for each $v_j \in \beta$, directly decompose $T(v_j) = \sum c_{ij} w_i$ as a linearly combination of $\gamma = \{ w_1, \dots, w_n \}$. The change of coordinate matrix is then formed by collecting the coefficients $([\mathrm{Id}]_\beta^\gamma)_{ij} = c_{ij}$
- find another ordered bases α, typically some *standard* basis[1], and compute the change of coordinate matrices $[\mathrm{Id}]_\beta^\alpha$ and $[\mathrm{Id}]_\gamma^\alpha$. The change of coordinate matrix $A = [\mathrm{Id}]_\beta^\gamma$ is then the matrix that solves $[\mathrm{Id}]_\gamma^\alpha A = [\mathrm{Id}]_\beta^\alpha$ (which can be solved with techniques from MATH1030, e.g. RREF), or more explicitly $A = \left([\mathrm{Id}]_\gamma^\alpha\right)^{-1}[\mathrm{Id}]_\beta^\alpha$

One reason to use standard basis is that the change of coordinate matrix is usually easy to find, e.g. on an ordered basis $\beta = \left\{ \begin{pmatrix} c_{11} \\ c_{21} \\ \vdots \\ c_{n1} \end{pmatrix}, \dots, \begin{pmatrix} c_{1n} \\ c_{2n} \\ \vdots \\ c_{nn} \end{pmatrix} \right\}$ of F^n, the change of coordinate matrix $[\mathrm{Id}]_\beta^\alpha$ to standard ordered

[1] For example, $\{ e_1, \dots, e_n \}$ for F^n, or $\{ 1, x, \dots, x^n \}$ for $\mathsf{P}_n(F)$.

basis $\alpha = \{ e_1, \ldots, e_n \}$ is just formed by collecting all entries into a matrix $[\text{Id}]_\beta^\alpha = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{n1} \\ c_{21} & \ldots & & \\ \vdots & & & \\ c_{1n} & \ldots & & c_{nn} \end{pmatrix}$, at the cost of the (potentially) more complicated $[\text{Id}]_\alpha^\beta = \left([\text{Id}]_\beta^\alpha\right)^{-1}$.

Similarly, if we want to find the matrix representation of a linear map $T \in L(V)$ with respect to some ordered basis β, we can

- for $v_j \in \beta$, directly decompose each of $T(v_j) = \sum_i c_{ij} v_i$ as a linear combination of β. The matrix representation $[T]_\beta$ is the formed by collecting the coefficients $([T]_\beta)_{ij} = c_{ij}$
- find another ordered basis α on which the matrix representation $[T]_\alpha$ is easier to find, then compute the change of coordinate matrices $[\text{Id}]_\alpha^\beta, [\text{Id}]_\beta^\alpha$. The matrix representation $[T]_\beta$ is then $[T]_\beta = [\text{Id}]_\alpha^\beta [T]_\alpha [\text{Id}]_\beta^\alpha$

In many cases, finding a basis α in which $[T]_\alpha$ has certain pattern (in *canonical form*) would give insights on how the linear map works, and (sometimes) ease the cost of computations.

1.2 Matrix Relations

Definition 1.2. Two matrices $A, B \in F^{n \times n}$ are *similar* if there exists an invertible matrix $Q \in F^{n \times n}$ such that $A = QBQ^{-1}$.

Definition 1.3. Two matrices $A, B \in F^{m \times n}$ are *equivalent* if there exist invertible matrices $P \in F^{m \times m}$ and $Q \in F^{n \times n}$ such that $A = PBQ^{-1}$.

Obviously, if β, γ are two ordered bases of V and $T \in L(V)$, $[T]_\beta$ and $[T]_\gamma$ must be similar. A similar statement holds for matrix equivalence as well. In exercise Q3 we will show the converse of this, that is two matrices are equivalent only if they represent the same linear map under different ordered bases. You can show a similar proposition on similar matrices as well.

2 Exercises

1. (Textbook Sec 2.5 Q13)

   Let V be a finite-dimensional vector space over a field F, and let $\beta = \{ x_1, \ldots, x_n \}$ be an ordered basis for V. Let Q be an $n \times n$ invertible matrix with entries from F. Define

   $$x'_j = \sum_{i=1}^{n} Q_{ij} x_i \quad \text{for } 1 \le j \le n,$$

   and set $\beta' = \{ x'_1, \ldots, x'_n \}$. Prove that β' is a basis for V and hence Q is the change of coordinate matrix changing β'-coordinates to β-coordinates.

   This shows in particular every invertible matrix is a change of coordinate matrix for some ordered bases.

   Solution: To show that β' is a basis for V, it suffices to show that it is linearly independent.
   Let $c_1, \ldots, c_n \in F$ such that $\sum_j c_j x'_j = 0$.
   By definition, $0 = \sum_j c_j x'_j = \sum_j \sum_i c_j Q_{ij} x_i = \sum_i (\sum_j Q_{ij} c_j) x_i$.
   Since β is linearly independent, we must have $\sum_j Q_{ij} c_j = 0$ for each i, or in matrix form

   $$Q \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

   As Q is invertible, the only possibility is that $c_1 = \ldots = c_n = 0$. This implies that β' is linearly independent.
   Thus β' is a basis for V.

To find the change of coordinate matrix, we compute the matrix representation $[\mathrm{Id}]_{\beta'}^{\beta}$.
For each $x'_j \in \beta'$, $\mathrm{Id}(x'_j) = x'_j = \sum_i Q_{ij} x_i$ with $x_i \in \beta$, which is a decomposition of $\mathrm{Id}(x'_j)$ in $\beta = \{ x_1, \ldots, x_n \}$.
So the change of coordinate matrix is $[\mathrm{Id}]_{\beta'}^{\beta} = Q$.

Note

Here we are actually proving something quite strong: given *any* basis β and *any* invertible matrix Q, we can always construct a new basis γ that the change of coordinate matrix $[\mathrm{Id}]_{\gamma}^{\beta}$ is exactly Q.

As we can see, although we transformed a basis β into a new basis β' using Q, the change of coordinate matrix from β-coordinates to β'-coordinates is $[\mathrm{Id}]_{\beta}^{\beta'} = Q^{-1}$, *not* Q. More precisely, the **coordinate** of a vector v changes according to Q^{-1} as $[v]_{\beta'} = [\mathrm{Id}]_{\beta}^{\beta'}[v]_\beta = Q^{-1}[v]_\beta$, *not* according to Q.
This can be understood in the following way: when we transform the scale we use, the *measurement* on the same object must also transform in an opposite way to counteract this change, for otherwise the result would be inconsistent.
Since this changes *against* the transformation of basis, the components (of the coordinate of a vector) are said to be *contravariant*. In some (usually old) references, this is (somewhat confusingly) shortened to "vectors transform contravariantly".

2. Let V be a finite dimensional vector space over F, and $\beta = \{ v_1, \ldots, v_n \}$ be an ordered basis of V (with $n = \dim(V)$). Let $Q \in F^{n\times n}$ be invertible, and for each j let $w_j = \sum_i Q_{ij} v_i$.

   From last question we know that $\gamma = \{ w_1, \ldots, w_n \}$ is also an ordered basis of V with $[\mathrm{Id}_V]_{\gamma}^{\beta} = Q$.

   Let $\beta^* = \{ f_1, \ldots, f_n \}$ and $\gamma^* = \{ g_1, \ldots, g_n \}$ be the (ordered) dual bases corresponding to $\{ v_1, \ldots, v_n \}$ and $\{ w_1, \ldots, w_n \}$ respectively. Find the change of coordinate matrix $[\mathrm{Id}_{V^*}]_{\gamma^*}^{\beta^*}$.

   (Recall that the dual basis $\{ f_1, \ldots, f_n \}$ corresponding to basis $\{ v_1, \ldots, v_n \}$ is a set of linear maps $f_i \in V^* = L(V, F)$ such that $f_i(v_j) = \delta_{ij}$ for each i, j, which forms a basis of V^*)

Solution: By definition, $P = [\mathrm{Id}_{V^*}]_{\gamma^*}^{\beta^*}$ satisfies $g_j = \mathrm{Id}_{V^*}(g_j) = \sum_k P_{kj} f_k$. So

$$\begin{aligned}\delta_{ij} = g_j(w_i) &= \sum_k P_{kj} f_k(w_i) \\ &= \sum_k \sum_l P_{kj} Q_{li} f_k(v_l) \\ &= \sum_k P_{kj} Q_{ki} \\ &= \sum_k (Q^\mathsf{T})_{ik} P_{kj}\end{aligned}$$

or in matrix notation

$$I_n = Q^\mathsf{T} P$$

This implies that $[\mathrm{Id}_{V^*}]_{\gamma^*}^{\beta^*} = P = \left(Q^\mathsf{T}\right)^{-1}$.

Note

It may be more natural to look at $[\mathrm{Id}_{V^*}]_{\beta^*}^{\gamma^*} = \left([\mathrm{Id}_{V^*}]_{\gamma^*}^{\beta^*}\right)^{-1} = Q^\mathsf{T}$.

If we write the coordinate of a vector in the dual space $f \in L(V, F) = V^*$ as a *row vector* instead of a column vector, this indicates that the **coordinate** of a dual vector changes according to Q as

$[f]_{\gamma^*} = [f]_{\beta^*}Q$ (with the matrix Q multiplied on the *right* instead of on the *left*). Since this changes *along with* the transformation of basis, the components (of the coordinate) are said to be *covariant*.

Noting the reversal of left/right multiplication above, together with the transpose map $t : L(V, W) \to L(W^*, V^*)$ mentioned in the last tutorial session, we can see a reoccurring theme:

The operation of "taking dual" "reverses" the "order"

I am using quotes because the meanings of these terms are not well-defined (yet), although they can be rigorously defined with some effort. Because of such reversal, this "dual operation" is *also* said to be *contravariant*. For those who are eager to know more (about this "dual operation"), you may want to check out the relevant sections about dual spaces in the reference books (both Friedberg et al. and Axler), or if you are feeling *particularly adventurous*, *Algebra* by Mac Lane and Birkhoff.

3. (See also textbook Sec. 2.5 Q14)

   Let V, W be finite dimensional vector spaces over F with dimension $\dim(V) = n$, $\dim(W) = m$, and $B, C \in F^{m\times n}$. Show that B, C are equivalent if and only if there exist $T \in L(V, W)$ and ordered bases β, β' of V and γ, γ' of W such that $[T]^\gamma_\beta = B$, $[T]^{\gamma'}_{\beta'} = C$.

   (In another word, two matrices of the same shape are equivalent if and only if they represent the same linear map under different ordered bases)

Solution: Suppose $B = [T]^\gamma_\beta$ and $C = [T]^{\gamma'}_{\beta'}$ for some $T, \beta, \gamma, \beta,', \gamma'$. Then $B = [T]^\gamma_\beta = [\mathrm{Id}_W \circ T \circ \mathrm{Id}_V]^\gamma_\beta = [\mathrm{Id}_W]^\gamma_{\gamma'}[T]^{\gamma'}_{\beta'}[\mathrm{Id}_V]^{\beta'}_\beta = ([\mathrm{Id}_W]^\gamma_{\gamma'})\, C \left([\mathrm{Id}_V]^\beta_{\beta'}\right)^{-1}$, so B, C are equivalent.

Suppose B, C are equivalent.
Then there exists invertible $P \in F^{m\times m}$, $Q \in F^{n\times n}$ such that $B = PCQ^{-1}$.
Let $\beta = \{ v_1, \dots, v_n \}$, $\gamma = \{ w_1, \dots, w_m \}$ be ordered bases of V and W respectively, and $T \in L(V, W)$ be the linear map that satisfies $T(v_j) = \sum_i B_{ij} w_i$ for all j.
By definition, $[T]^\gamma_\beta = B$.

By Question 1, there exist ordered bases β' of V and γ' of W such that $Q = [\mathrm{Id}_V]^\beta_{\beta'}$ is the change of coordinate matrix from β'-coordinates to β-coordinates, and $P = [\mathrm{Id}_W]^\gamma_{\gamma'}$ is the change of coordinate matrix from γ'-coordinates to γ-coordinates.
Then $C = P^{-1}BQ = \left([\mathrm{Id}_W]^\gamma_{\gamma'}\right)^{-1}[T]^\gamma_\beta[\mathrm{Id}_V]^\beta_{\beta'} = [\mathrm{Id}_W]^{\gamma'}_\gamma[T]^\gamma_\beta[\mathrm{Id}_V]^\beta_{\beta'} = [T]^{\gamma'}_{\beta'}$.

4. Show that for every matrix $A \in F^{m\times n}$, there exists a unique $r \in \{ 0, \dots, \min(m, n) \}$ such that A is equivalent to the matrix

$$S_r = \begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix}$$

Solution: Let $r = \mathrm{rank}(L_A) \in \{ 0, \dots, \min(m, n) \}$ be the (column) rank of A. We will show that A is equivalent to S_r.

Let $\{ y_1, \dots, y_r \}$ be a basis of $\mathsf{R}(L_A) \subseteq F^m$, and extend it to a basis $\gamma = \{ y_1, \dots, y_m \}$ of F^m. For $i \leq r$ let $x_i \in F^n$ such that $Ax_i = y_i$.
Since $\{ y_1, \dots, y_r \}$ is linearly independent, so is $\{ x_1, \dots, x_r \}$, and we can extend it to a basis $\beta = \{ x_1, \dots, x_n \}$ of F^n.

For $i \leq r$, $Tx_i = y_i$, so $[Tx_i]_\gamma = e_i$.
For $i > r$, as $\{ y_1, \dots, y_r \} = \{ Tx_1, \dots, Tx_r \}$ spans $\mathsf{R}(L_A)$, we must have $Tx_i = 0$, so $[Tx_i]_\gamma = 0$.
By definition, $[T]^\gamma_\beta = \begin{pmatrix} I_{r\times r} & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{pmatrix} = S_r$, so on the standard bases α_n, α_m of F_n, F_m

respectively, $A = [L_A]_{\alpha_n}^{\alpha_m} = [\mathrm{Id}_{F^m}]_{\gamma}^{\alpha_m}[L_A]_{\beta}^{\gamma}[\mathrm{Id}_{F^n}]_{\alpha_n}^{\beta} = [\mathrm{Id}_{F^m}]_{\gamma}^{\alpha_m} S_r \left([\mathrm{Id}_{F^n}]_{\beta}^{\alpha_n}\right)^{-1}$ which is equivalent to S_r.

We now show the uniqueness of such r.
Let $r' \in \{ 0, \ldots, \min(m, n) \}$ such that A is also equivalent to $S_{r'}$. Then S_r and $S_{r'}$ are equivalent.
By Question 3, there exist a linear map $T \in L(V, W)$ between finite dimensional vector spaces and ordered bases β, β' of V, γ, γ' of W such that $[T]_{\beta}^{\gamma} = S_r$ and $[T]_{\beta'}^{\gamma'} = S_{r'}$.
WLOG let elements of β, γ be $\beta' = \{ v_1, \ldots, v_n \}$ and $\gamma = \{ w_1, \ldots, w_m \}$.
By construction, for $i \leq r$, $[Tv_i]_{\gamma} = e_i$, so $w_i = Tv_i \in \mathsf{R}(T)$. As $\{ w_1, \ldots, w_r \}$ is linearly independent, $\mathrm{rank}(T) \geq r$.
Similarly, for $i > r$, $[Tv_i]_{\gamma} = 0$, so $Tv_i = 0$, $v_i \in \mathsf{N}(T)$. Since $\{ v_{r+1}, \ldots, v_n \}$ is linearly independent, $\mathrm{nullity}(T) \geq n - r$.
By dimension theorem, $n = \dim(V) = \mathrm{rank}(T) + \mathrm{nullity}(T) \geq r + (n - r) = n$, so we must have $\mathrm{rank}(T) = r$. Repeat the same argument on β' and γ' we obtain $\mathrm{rank}(T) = r'$. So $r = r'$.
Thus such r is unique.

Note

In another word, such S_r is a *canonical form* when we are concerning with matrix equivalence (as defined above), and the rank $r = \mathrm{rank}(L_A) = \mathrm{rank}(A)$ uniquely determines the matrices that A is equivalent to.

In the proof of uniqueness, more canonically we can take $V = F^n$ and $W = F^m$, and the corresponding bases the standard ordered bases. This should make the constructions more explicit.