HW 6 Dual Stuff

Liu Mingdao 2020011156

Exercise 1.6.1

1. We know

$$L([1,\cdots,x^{n-1}]) = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} \begin{bmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{bmatrix}$$

Then the matrix of L is

$$\begin{bmatrix} 1 & a_1 & \cdots & a_1^{n-1} \\ 1 & a_2 & \cdots & a_2^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & a_n & \cdots & a_n^{n-1} \end{bmatrix}$$

It is a Vandermonde matrix.

2. L is invertible $\iff$ the matrix of L is invertible. Note that for Vandermonde matrix L,

$$det(L) = \prod_{i>j,\ i,j\in\{1,\cdots,n\}} (a_i - a_j)$$

Then we have $a_1,\cdots,a_n$ are distinct $\iff$ $det(L)\neq 0$ $\iff$ L is invertible.

3. First prove that $\{ev_{a_1},\cdots,ev_{a_n}\}$ form a basis implies $a_1,\cdots,a_n$ are distinct. Suppose $\{ev_{a_1},\cdots,ev_{a_n}\}$ forms a basis and $\exists m,n\in\{1,\cdots,n\}$ s.t. $a_m = a_n (m\neq n)$. We know $ev_m = ev_n$, then $\{ev_{a_1},\cdots,ev_{a_n}\}$ are linearly dependent. It is contradictory to the assumption that $\{ev_{a_1},\cdots,ev_{a_n}\}$ is a basis of an n dimensional vector space V^*.

Then prove that $a_1,\cdots,a_n$ are distinct implies $\{ev_{a_1},\cdots,ev_{a_n}\}$ form a basis of V^*. From the previous subproblem, we know that there is a unique solution to the interpolation problem with n distinct points. So let $l_i = \prod\limits_{j\neq i}\frac{x-a_j}{a_i-a_j}$, $\{l_i\}$ is the Lagrange basis for V. For $\forall \boldsymbol{v}\in V$, we have

$$\boldsymbol{v}(x) = \sum_{i=1}^{n} ev_{a_i}(\boldsymbol{v})l_i$$

For $\forall\alpha\in V^*$, $\forall\boldsymbol{v}\in V$, we get

$$\alpha(\boldsymbol{v}) = \alpha(\sum_{i=1}^{n} ev_{a_i}(\boldsymbol{v})l_i) = \sum_{i=1}^{n}\alpha(l_i)ev_{a_i}(\boldsymbol{v})$$

So for any given distinct $a_1,\cdots,a_n$ and given $\alpha\in V^*$, α is a linear combination of $\{ev_{a_1},\cdots,ev_{a_n}\}$. I.e. $\{ev_{a_1},\cdots,ev_{a_n}\}$ forms a basis of V^*.

4. Consider polynomials in V. Let

$$[p_{-1},p_0,p_1] = [1,x,x^2]W$$

for some $W\in M_3(\mathbb{R})$. The dual vectors of $[p_{-1},p_0,p_1]$ are $[ev_{-1},ev_0,ev_1]$. Then we know

$$IW^{-1} = \begin{bmatrix} ev_{-1}(1) & ev_{-1}(x) & ev_{-1}(x^2) \\ ev_0(1) & ev_0(x) & ev_0(x^2) \\ ev_1(1) & ev_1(x) & ev_1(x^2) \end{bmatrix} = \begin{bmatrix} 1 & -1 & 1 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix}$$

So $W = \begin{bmatrix} 0 & 1 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} \\ \frac{1}{2} & -1 & \frac{1}{2} \end{bmatrix}$, we have $p_{-1} = \frac{1}{2}x^2 - \frac{1}{2}x,\ p_0 = 1-x^2,\ p_1 = \frac{1}{2}x^2+\frac{1}{2}x$.

5. Since $-1,\ 0,\ 1,\ 2$ are distinct, from subproblem 3 we know $\{ev_{-1}, ev_0, ev_1, ev_2\}$ forms a basis of $V*$. For $i \in -1, 0, 1, 2$, let $k_i = \prod\limits_{j \neq i} \frac{x-j}{i-j}$, then we have

$$ev_{-2} = \sum_{i=-1}^{2} ev_{-2}(k_i) \cdot ev_i$$

Specifically, we have

$$-ev_{-2} + 4ev_{-1} - 6ev_0 + 4ev_1 - ev_2 = 0$$

□

Exercise 1.6.2 Let $\boldsymbol{v}$ and $\boldsymbol{w}$ be two polynomials in V. To say α is a dual vector of V is equivalent to say that α is a linear map from V to $\mathbb{R}$. I.e. we only need to show that for $\forall \lambda,\ \mu \in \mathbb{R}$, $\alpha(\lambda \boldsymbol{v} + \mu \boldsymbol{w}) = \lambda \alpha(\boldsymbol{v}) + \mu \alpha(\boldsymbol{w})$. We write the 5 maps mentioned in this questions as $\alpha_1, \cdots, \alpha_5$.

1. α_1 is a dual vector. $\alpha_1(\lambda \boldsymbol{v} + \mu \boldsymbol{w}) = 6(\lambda \boldsymbol{v} + \mu \boldsymbol{w})(5) = \lambda \cdot 6\boldsymbol{v}(5) + \mu \cdot 6\boldsymbol{w}(5) = \lambda \alpha_1(\boldsymbol{v}) + \mu \alpha_1(\boldsymbol{w})$. So $\alpha_1 \in V^*$.

2. α_2 is not a dual vector. Consider $\boldsymbol{v} = x^2$. $\alpha_2(\boldsymbol{v}) = +\infty \notin \mathbb{R}$.

3. α_3 is a dual vector. This map takes the coefficient of term x^2 for a polynomial in V, so it is linear. Let $\boldsymbol{v} = a_0 + a_1 x + a_2 x^2$, $\boldsymbol{w} = b_0 + b_1 x + b_2 x^2$. We have $\alpha_3(\lambda \boldsymbol{v} + \mu \boldsymbol{w}) = \lim\limits_{x \to +\infty} \frac{\lambda(a_0 + a_1 x) + \mu(b_0 + b_1 x)}{x^2} + \lim\limits_{x \to +\infty} \frac{\lambda a_2 x^2 + \mu b_2 x^2}{x^2} = \lambda a_2 + \mu b_2 = \lambda \alpha_3(\boldsymbol{v}) + \mu \alpha_3(\boldsymbol{w})$.

4. α_4 is not a dual vector. For $\forall \boldsymbol{v} \in V$, let $\boldsymbol{v} = a_0 + a_1 x + a_2 x^2$, we have $\alpha_4(\boldsymbol{v}) = (8a_2 + a_1)(9a_2 + 3a_1 + a_0)$. It is not linear. Specifically, we show a conter example. Let $\boldsymbol{v} = 1,\ \boldsymbol{w} = x$, then $\alpha_4(\boldsymbol{v}) = 0, \alpha_4(\boldsymbol{w}) = 3, \alpha_4(\boldsymbol{v} + \boldsymbol{w}) = 4$, we get $\alpha_4(\boldsymbol{v}) + \alpha_4(\boldsymbol{w}) \neq \alpha_4(\boldsymbol{v} + \boldsymbol{w})$.

5. α_5 is not a dual vector, since the degree of a polynomial is not a linear function of the polynomial. Specifically, let $\boldsymbol{v} = \boldsymbol{w} = x$, then $\boldsymbol{v} + \boldsymbol{w} = 2x$, $deg(\boldsymbol{v}) = deg(\boldsymbol{w}) = 1,\ deg(\boldsymbol{v} + \boldsymbol{w}) = 1 \neq deg(\boldsymbol{v}) + deg(\boldsymbol{w})$. □

Exercise 1.6.3 Let $\boldsymbol{v} = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. From the differentiability of f, we have $f(\boldsymbol{p} + t\boldsymbol{v}) - f(\boldsymbol{p}) = tv_1 \frac{\partial f}{\partial x_1}(\boldsymbol{p}) + tv_2 \frac{\partial f}{\partial x_2}(\boldsymbol{p}) + o(||v||t)$, then $\nabla_{\boldsymbol{v}} f = \lim\limits_{t \to 0}(v_1 \frac{\partial f}{\partial x_1}(\boldsymbol{p}) + v_2 \frac{\partial f}{\partial x_2}(\boldsymbol{p}) + \frac{o(||v||t)}{t}) = \begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{p}) & \frac{\partial f}{\partial x_2}(\boldsymbol{p}) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}$. So $\nabla_{\boldsymbol{v}} f$ is a linear map from $\mathbb{R}^2$ to $\mathbb{R}$, and is thus a dual vector in $(\mathbb{R}^2)^*$. Its "coordinates" under standard dual basis are $\begin{bmatrix} \frac{\partial f}{\partial x_1}(\boldsymbol{p}) & \frac{\partial f}{\partial x_2}(\boldsymbol{p}) \end{bmatrix}$. (Of course, the dual basis is $\{\begin{bmatrix} 1 & 0 \end{bmatrix}, \begin{bmatrix} 0 & 1 \end{bmatrix}\}$) □

Exercise 1.6.4

1. For $\forall \alpha \in Ker(L^*),\ L^*(\alpha) = \boldsymbol{0}$. Then we know for $\forall \boldsymbol{v} \in V,\ L^*(\alpha)(v) = \alpha \circ L(v) = \boldsymbol{0}$. For $\forall \boldsymbol{w} \in Ran(L)$, we have $\boldsymbol{w} = L(\boldsymbol{v})$ for some $\boldsymbol{v} \in V$. So $\alpha(\boldsymbol{w}) = \alpha \circ L(v) = \boldsymbol{0}$.
Conversely, suppose $\forall \alpha \in W^*$ s.t. for $\forall \boldsymbol{w} \in Ran(L),\ \alpha(\boldsymbol{w}) = \boldsymbol{0}$. Since for $\forall \boldsymbol{v} \in V,\ L(\boldsymbol{v}) \in Ran(L)$, we have $\alpha \circ L(v) = \alpha(L(v)) = \boldsymbol{0}$. Note that $\boldsymbol{v}$ is arbitrary, so $L^*(\alpha) = \alpha \circ L = \boldsymbol{0}$. I.e. $\alpha \in Ker(L^*)$.

2. For $\forall \beta \in Ran(L^*)$, we have $\beta = L^*(\gamma) = \gamma \circ L$ for some $\gamma \in W^*$. Then for $\forall \boldsymbol{v} \in Ker(L), \beta(\boldsymbol{v}) = \gamma(L(\boldsymbol{v})) = \boldsymbol{0}$.
Conversely, we try to prove if for $\forall \alpha \in V^*$ s.t. $\forall z \in Ker(L),\ \alpha(z) = \boldsymbol{0}$, then $\alpha \in Ran(L^*)$.

We first show that for $\forall \boldsymbol{v}, \boldsymbol{w} \in V$ s.t. $L(\boldsymbol{v}) = L(\boldsymbol{w})$, we have $\alpha(\boldsymbol{v}) = \alpha(\boldsymbol{w})$. **Proof.** If $L(\boldsymbol{v}) = L(\boldsymbol{w})$, then $L(\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{0}$. I.e. $(\boldsymbol{v} - \boldsymbol{w}) \in Ker(L)$, so $\alpha(\boldsymbol{v} - \boldsymbol{w}) = \boldsymbol{0}$, i.e. $\alpha(\boldsymbol{v}) = \alpha(\boldsymbol{w})$. Thus, we know that the value of $\alpha(\boldsymbol{v})$ is only dependent on $L(\boldsymbol{v})$.

Suppose $dim\ W = m$. We pick a basis for $Ran(L)$, i.e. $Ran(L) = span(\{w_1, \cdots, w_k\})$. Then expand this basis such that $W = span(\{w_1, \cdots, w_m\})$. Let $W' = span(\{w_{k+1}, \cdots, w_m\})$, we have $W = Ran(L) \oplus W'$, i.e. $\forall \boldsymbol{x} \in W$, there is a unique decomposition $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2$ s.t. $\boldsymbol{x}_1 \in Ran(L),\ \boldsymbol{x}_2 \in W'$. Then we define a map $\beta : W \to \mathbb{R}$, s.t. $\beta(\boldsymbol{x}) = \alpha(\boldsymbol{v})$, where $\boldsymbol{v}$ is any preimage of $\boldsymbol{x}_1$ (We define $\boldsymbol{x}_1$ as such: $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2,\ \boldsymbol{x}_1 \in Ran(L),\ \boldsymbol{x}_2 \in W'$). We know β is well defined, since the value of $\alpha(\boldsymbol{v})$ depends only on $\boldsymbol{x}_1 = L(\boldsymbol{v})$. We immdeiately know that $\beta(\boldsymbol{x}) = \beta(\boldsymbol{x}_1)$.

Then we try to show that β is a linear map. Suppose $\forall \boldsymbol{x}, \boldsymbol{y} \in W$, we decompose them into $\boldsymbol{x} = \boldsymbol{x}_1 + \boldsymbol{x}_2,\ \boldsymbol{x}_1 \in Ran(L),\ \boldsymbol{x}_2 \in W'$ and $\boldsymbol{y} = \boldsymbol{y}_1 + \boldsymbol{y}_2,\ \boldsymbol{y}_1 \in Ran(L),\ \boldsymbol{y}_2 \in W'$. Let $\boldsymbol{v}, \boldsymbol{w} \in V$ be any vector s.t. $\boldsymbol{x}_1 = L(\boldsymbol{v}),\ \boldsymbol{y}_1 = L(\boldsymbol{w})$. Note that $\boldsymbol{x}_1 + \boldsymbol{y}_1 = L(\boldsymbol{v} + \boldsymbol{w}), k\boldsymbol{x}_1 = L(k\boldsymbol{v})$, we have $\beta(\boldsymbol{x} + \boldsymbol{y}) = \alpha(\boldsymbol{v} + \boldsymbol{w}) = \alpha(\boldsymbol{v}) + \alpha(\boldsymbol{w}) = \beta(\boldsymbol{x}_1) + \beta(\boldsymbol{y}_1) = \beta(\boldsymbol{x}) + \beta(\boldsymbol{y}),\ \beta(k\boldsymbol{x}) = \alpha(k\boldsymbol{v}) = k\alpha(\boldsymbol{v}) = k\beta(\boldsymbol{x}_1) = k\beta(\boldsymbol{x})$. So β is a linear map, i.e. $\beta \in W^*$.

Then $\forall \boldsymbol{v} \in V$, $\alpha(\boldsymbol{v}) = \beta(L(\boldsymbol{v})) = L^*(\beta)(\boldsymbol{v})$, i.e. $\alpha \in Ran(L^*)$.

□

Exercise 1.6.5

1. Bilinear:

$$\begin{aligned}
(\lambda \boldsymbol{a} + \mu \boldsymbol{b}, \boldsymbol{w}) &= (\lambda \boldsymbol{a} + \mu \boldsymbol{b})^T A \boldsymbol{w} \\
&= \lambda \boldsymbol{a}^T A \boldsymbol{w} + \mu \boldsymbol{b}^T A \boldsymbol{w} \\
&= \lambda(\boldsymbol{a}, \boldsymbol{w}) + \mu(\boldsymbol{b}, \boldsymbol{w}) \\
(\boldsymbol{v}, \lambda \boldsymbol{a} + \mu \boldsymbol{b}) &= \boldsymbol{v}^T A(\lambda \boldsymbol{a} + \mu \boldsymbol{b}) \\
&= \lambda \boldsymbol{v}^T A \boldsymbol{a} + \mu \boldsymbol{v}^T A \boldsymbol{b} \\
&= \lambda(\boldsymbol{v}, \boldsymbol{a}) + \mu(\boldsymbol{v}, \boldsymbol{b})
\end{aligned}$$

Symetirc: Note that A is symetirc, we have $(\boldsymbol{v}, \boldsymbol{w}) = (\boldsymbol{v}, \boldsymbol{w})^T = \boldsymbol{w}^T A^T \boldsymbol{v} = \boldsymbol{w}^T A \boldsymbol{v} = (\boldsymbol{w}, \boldsymbol{v})$.

Positive definite: Note that A is positive definite, we know

$$\begin{cases} \boldsymbol{v}^T A \boldsymbol{v} > 0 \iff v \neq 0 \\ \boldsymbol{v}^T A \boldsymbol{v} = 0 \iff v = 0 \end{cases}$$

So we have $(\boldsymbol{v}, \boldsymbol{v}) \geq 0$ and $(\boldsymbol{v}, \boldsymbol{v}) = 0 \iff \boldsymbol{v} = \boldsymbol{0}$.

2. Note that $(A^{-1}\boldsymbol{v}, \boldsymbol{w}) = \boldsymbol{v}^T (A^{-1})^T A^T \boldsymbol{w} = \boldsymbol{v}^T (A^T)^{-1} A \boldsymbol{w} = \boldsymbol{v}^T \boldsymbol{w}$, we have $Riesz(\boldsymbol{v}^T) = A^{-1}\boldsymbol{v}$.

3. We know $Reisz^{-1}(\boldsymbol{v}) = (\boldsymbol{v}|$, so $Reisz^{-1}(\boldsymbol{v}) = \boldsymbol{v}^T A$.

4. First, we prove

Lemma For any linear bijection $A : V \to W$, $(A^*)^{-1} = (A^{-1})^*$.

Proof. Suppose arbitrary $\boldsymbol{v} \in V, \boldsymbol{w} \in W, \boldsymbol{\alpha} \in V^*, \boldsymbol{\beta} \in W^*$ s.t. $\boldsymbol{w} = A(\boldsymbol{v}), \boldsymbol{\alpha} = A^*(\boldsymbol{\beta})$. We have $(A^{-1})^*(\boldsymbol{\alpha})(\boldsymbol{w}) = \boldsymbol{\alpha} \circ A^{-1}(\boldsymbol{w}) = \boldsymbol{\alpha}(\boldsymbol{v}) = A^*(\boldsymbol{\beta})(\boldsymbol{v}) = \boldsymbol{\beta} \circ A(\boldsymbol{v}) = \boldsymbol{\beta}(\boldsymbol{w}) = (A^*)^{-1}(\boldsymbol{\alpha})(\boldsymbol{w})$. Thus, $(A^*)^{-1} = (A^{-1})^*$.

Then, let L be the bra map $L : \mathbb{R}^n \to (\mathbb{R}^n)^*$, $\boldsymbol{v}, \boldsymbol{w} \in \mathbb{R}^n$. Consider $L^* : (\mathbb{R}^n)^{**} \to (\mathbb{R}^n)^*$, we have $L^*(ev_{\boldsymbol{v}})(\boldsymbol{w}) = ev_{\boldsymbol{v}} \circ L(\boldsymbol{w}) = ev_{\boldsymbol{v}}(\ (\boldsymbol{w}|\ ) = (\boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{v}, \boldsymbol{w}) = L(\boldsymbol{v})(\boldsymbol{w})$. With the canonical isomorphism between $(\mathbb{R}^n)^{**}$ and $\mathbb{R}^n$, we can say $L = L^*$. Note that $Reisz = L^{-1}$, from **Lemma**, we know $Reisz^* = Reisz$. So $Reisz^*(\boldsymbol{v}^T) = Reisz(\boldsymbol{v}^T) = A^{-1}\boldsymbol{v}$.

5. From the previous subproblem, we know $Reisz^* = Reisz$, then $Reisz^{ad} = Reisz^{-1} \circ Reisz^* \circ Reisz^{-1} = Reisz^{-1} \circ Reisz \circ Reisz^{-1} = Reisz^{-1}$. We have $Reisz^{ad}(\boldsymbol{v}) = Reisz^{-1}(\boldsymbol{v}) = \boldsymbol{v}^T A$. □