
Linear Vector Spaces and Duals

This chapter is devoted to some fundamentals of linear vector spaces. First, we define the linear vector space.

Definition 1 (Linear vector space). *A set V with well defined addition and scalar multiplication, i.e. $\forall u, v \in V$ and $\alpha \in \mathbb{R}$, there is $w = u + v \in V$ and $x = \alpha u \in V$, is called linear vector space if the following equations hold for any $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$*

$$(u + v) + w = u + (v + w)$$

$$u + v = v + u$$

$$\exists \mathbf{0} \in V, \text{ s.t. } u + \mathbf{0} = u$$

$$\exists -u \in V, \text{ s.t. } u + (-u) = 0$$

$$(\alpha + \beta)u = \alpha u + \beta u$$

$$(\alpha\beta)u = \alpha(\beta u)$$

$$\alpha(u + v) = \alpha u + \alpha v$$

$$1u = u.$$

Given a linear vector space V , for scalars $\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n \in \mathbb{R}$ and vectors $v_1, v_2, v_3, \dots, v_n \in V$, we say they are linearly independent if the equation

$$(1.1) \quad \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 + \dots + \alpha_n v_n = 0$$

can only be satisfied when $\alpha_i = 0$ for $i = 1, 2, \dots, n$. We say that $v_1, v_2, v_3, \dots, v_n$ is a basis if they are linearly independent and an arbitrary vector $v \in V$ can be expressed as the linear combination of $v_1, v_2, v_3, \dots, v_n$.

The dimension of a vector space is the number of vectors in any basis for the space.

Example 1. \mathbb{R}^n is a well known example of linear vector space. A standard basis of

this is $e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \\ 0 \end{pmatrix}, \dots, e_n = \begin{pmatrix} 0 \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$. Here there are n components for the vector e_i , and 1 is the i -th component.

1.1 Norms and inner products

Definition 2 (Norm). A norm on the vector space V is a function $\|\cdot\| : V \mapsto \mathbb{R}$ that satisfies, for any $u, v \in V$ and $\alpha \in \mathbb{R}$, the following three properties

1. $\|v\| \geq 0$ and $\|v\| = 0$ iff $v = 0$;
2. $\|\alpha v\| = |\alpha| \|v\|$;
3. $\|u + v\| \leq \|u\| + \|v\|$.

Example 2. $V = \mathbb{R}^n$ and $p \geq 1$. Define

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}.$$

It can be verified that $\|\cdot\|$ satisfies all the three conditions in the above definition and hence it defines a norm in \mathbb{R}^n . Of the most important are $p = 1, 2$ and ∞ :

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \quad \text{and} \quad \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|.$$

Theorem 1. Assume that V is a finite dimensional vector space and $\|\cdot\|$ and $\|\cdot\|_*$ are given two norms on V . Then there exist two positive constants c and C such that

$$c\|v\| \leq \|v\|_* \leq C\|v\|.$$

Proof. Assume that v_1, v_2, \dots, v_n is a basis of V . We define a function on \mathbb{R}^n by $f(x) = \|v\|_*$ for $x \in \mathbb{R}^n$ such that $v = \sum_{i=1}^n x_i v_i$. It is straightforward to show that f is a positive, continuous function on the compact set $S = \{x \in \mathbb{R}^n, \|v\| = 1\}$. By the well-known property of continuous function, we have

$$0 < c = \min_{x \in S} f(x), \quad C = \max_{x \in S} f(x) < \infty.$$

Therefore $c \leq \|v\|_* \leq C$ if $\|v\| = 1$. Applying the above inequality with $v/\|v\|$ for general $v \neq 0$ then leads to the desired inequality. \square

Operator norms

Assume that V and W are two linear vector spaces over \mathbb{R} or \mathbb{C} . Recall that a mapping¹ $A : V \mapsto W$ is said to be linear if

$$A(\alpha u + \beta v) = \alpha Au + \beta Av \quad \forall u, v \in V, \alpha, \beta \in \mathbb{R}.$$

The set of all linear operators from V to W will be denoted by $L(V, W)$ and $L(V) = L(V, V)$. Especially, $\mathcal{L}(V)$ denote the space of linear operators from \mathbb{V} to \mathbb{R} . Notice that $\mathcal{L}(V)$ is a linear vector space, hence a norm of an operator $A \in L(\mathbb{V})$ can be defined by the Definition (2) in general.

Example 3. For a matrix $A \in \mathbb{R}^{n \times m}$, we can treat A as a vector of size nm and define the following “entrywise” norm:

$$\|A\|_p = \left(\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^p \right)^{1/p}.$$

For $p = 2$, this is called the Frobenious norm:

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^m |a_{ij}|^2 \right)^{1/2}.$$

For $p = \infty$, this is called the max norm:

$$\|A\|_{\max} = \max\{|a_{ij}|\}.$$

Of the most useful is to define an operator norm using a vector norm.

Proposition 1. Assume that $\|\cdot\|$ is a norm on V . Then

$$\|A\| = \sup_{v \in V} \frac{\|Av\|}{\|v\|} \quad \forall A \in L(V)$$

defines a norm on $L(V)$ that satisfies, for any $A, B \in L(V), v \in V$

$$\|Av\| \leq \|A\|\|v\|$$

and

$$(1.2) \quad \|AB\| \leq \|A\|\|B\|.$$

An operator norm $\|\cdot\|$ is said to be *submultiplicative* if it satisfies (1.2). A vector norm $\|\cdot\|_\alpha$ and an operator norm $\|\cdot\|_\beta$ are said to be consistent if

$$\|Av\|_\alpha \leq \|A\|_\beta \|v\|_\alpha \quad \forall A \in L(V), v \in V.$$

¹ Mapping, transformation and operator are synonyms in this book.

Example 4. For the vector norms defined in (2), the corresponding matrix norms are given by

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|, \quad \|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|.$$

The matrix norm $\|A\|_2$ will be given later.

We shall now give the definitions of inner products.

Definition 3 (Inner Product). An inner product of V is a bilinear form $(\cdot, \cdot) : V \times V \mapsto \mathbb{R}$ such that for any $u, v, w \in V$ and $\alpha, \beta \in \mathbb{R}$.

1. $(v, v) \geq 0$ and $(v, v) = 0$ iff $v = 0$;
2. $(u, v) = (v, u)$;
3. $(\alpha u + \beta v, w) = \alpha(u, w) + \beta(v, w)$

Example 5. For $V = \mathbb{R}^n$. The Euclidean inner product is defined by

$$(x, y)_2 = \sum_{i=1}^n x_i y_i.$$

Lemma 1 (Cauchy inequality). Assume that (\cdot, \cdot) is an inner product on V , then

$$|(u, v)| \leq \|u\| \|v\| \quad \forall u, v \in V$$

where $\|v\| = (v, v)^{\frac{1}{2}}$ is a norm (induced by (\cdot, \cdot)) on V .

Proof. By the definition of inner product, we have, for any $u, v \in V$ and $t \in \mathbb{R}$

$$\|u\|^2 + t^2 \|v\|^2 - 2t(u, v) = (u - tv, u - tv) \geq 0.$$

Thus the desired inequality follows by taking $t = \|u\|/\|v\|$ (if $v \neq 0$). Moreover

$$\begin{aligned} \|u + v\|^2 &= (u + v, u + v) = \|u\|^2 + \|v\|^2 + 2(u, v) \\ &\leq \|u\|^2 + \|v\|^2 + 2\|u\|\|v\| \leq (\|u\| + \|v\|)^2. \end{aligned}$$

Thus $\|\cdot\|$ is indeed a norm. \square

Example 6. $V = \mathbb{R}^n$. For the inner product given by (5), the Cauchy-Schwarz inequality gives

$$\left| \sum_{i=1}^n x_i y_i \right|^2 \leq \sum_{i=1}^n |x_i|^2 \sum_{i=1}^n |y_i|^2.$$

Example 7. $V = L^2(\Omega)$, the vector space consisting of square integrable functions over in a domain $\Omega \subset \mathbb{R}^d$. Define an inner product by

$$(f, g)_{L^2} = \int_{\Omega} f(x)g(x)dx.$$

The corresponding Cauchy-Schwarz inequality is

$$\left| \int_{\Omega} f(x)g(x)dx \right|^2 \leq \int_{\Omega} |f(x)|^2 dx \int_{\Omega} |g(x)|^2 dx.$$

1.2 Examples of vector spaces

1.2.1 Polynomial space

Given a domain $\Omega \subset \mathbb{R}^d$ with $d \geq 1$, we consider the space of polynomials of two different types.

The first is the polynomial of degree not greater than $m \geq 0$:

$$(1.3) \quad \mathbb{P}_m(\Omega) = \left\{ p = \sum_{|\alpha| \leq m} a_\alpha x_1^{\alpha_1} \dots x_d^{\alpha_d} \right\}$$

where $\alpha = (\alpha_1, \dots, \alpha_d)$ and $|\alpha| = \sum_{i=1}^d \alpha_i$. We have

$$(1.4) \quad \dim \mathbb{P}_m(\Omega) = \binom{d+m}{m}$$

The second is the polynomial whose degree is not greater than $m \geq 0$ in each variable:

$$(1.5) \quad \mathbb{Q}_m(\Omega) = \left\{ p = \sum_{|\alpha| \leq m} a_\alpha x_1^{\alpha_1} \dots x_d^{\alpha_d} \right\}$$

$$(1.6) \quad \dim \mathbb{Q}_m(\Omega) = (m+1)^d$$

We are especially interested in polynomials in space dimension $d = 1, 2, 3$ for

$$\mathbb{P}_m(S_d) \text{ and } \mathbb{P}_m(C_d).$$

Here S_d is a d -dimensional simplex. For example, a triangle is a two dimensional simplex, and a tetrahedron is a three dimensional simplex. C_d is a d -dimensional cube. For example, $C_d = (0, 1)^d$.

1.2.2 Space of polynomial satisfying boundary conditions

Associated with the interval $(0, 1)$, we can also define the following linear vector space of dimension n :

$$(1.7) \quad V_n = \left\{ p : p \in \mathbb{P}_{n+2}(0, 1), p(0) = p(1) = 0. \right\}$$

This space also has a natural basis known as the nodal basis, which consists of the unique set of functions $\{\phi_i\} \subset V_n$ satisfying

$$\phi_i(x_j) = \delta_{ij}$$

which is given by, for $i = 1, 2, \dots, n$

$$(1.8) \quad \phi_i(x) = \frac{\prod_{j \neq i} (x - x_j)}{\prod_{j \neq i} (x_i - x_j)}$$

1.3 Eigenvalues and eigenvectors

We shall review the concept of eigenvalue and eigenvectors. An eigenvalue of the operator $A \in L(V)$ is a real or complex number λ which, for some nonzero vector $v \in V$, satisfies

$$(1.9) \quad (A - \lambda I)v = 0.$$

Any nonzero vector v satisfying the above equation is called an eigenvector of A corresponding to the eigenvalue λ .

If $A \in \mathbb{R}^{n \times n}$, obviously λ is an eigenvalue of A iff

$$\det(A - \lambda I) = 0.$$

Given $A \in L(V)$, we shall use the notation $\sigma(A)$ to denote the set of all eigenvalues of A . The spectral radius of A is defined and denoted by

$$\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|.$$

Proposition 2. For any $A \in L(V, \mathbb{W})$, $B \in L(\mathbb{W}, V)$, $\sigma(AB) \setminus \{0\} = \sigma(BA) \setminus \{0\}$ and $\sigma(A) = \sigma(T^{-1}AT)$ for any invertible $T \in L(V)$.

Proof. Assume $\lambda \in \sigma(BA) \setminus \{0\}$, then there exists $c \in V \setminus \{0\}$ such that

$$BAc = \lambda c.$$

This implies $Ac \neq 0$ and

$$AB(Ac) = \lambda Ac$$

which implies $\lambda \in \sigma(AB) \setminus \{0\}$.

The desired result then follows easily. \square

Proposition 3. Assume $\|\cdot\|$ is a submultiplicative norm on $L(V)$. Then

$$\rho(A) \leq \|A\|, \quad \forall A \in L(V).$$

Proposition 4. Assume that $A \in L(V)$, then

$$\lim_{k \rightarrow \infty} A^k = 0 \quad \text{if and only if } \rho(A) < 1.$$

1.4 Spectrum of symmetric and SPD matrices

We begin our discussion in the Euclidean space \mathbb{R}^n . An operator on \mathbb{R}^n is a matrix, say $A = (a_{ij}) \in \mathbb{R}^{n \times n}$. When A is said to be self-adjoint or symmetric, it often means that $A = A^t$ (here “ t ” denotes the transposition), that is, $a_{ij} = a_{ji}$, or equivalently

$$(Ax, y)_\ell = (x, Ay)_\ell \quad \forall x, y \in \mathbb{R}^n,$$

where $(\cdot, \cdot)_\ell$ is given in Example 5. In general, we have

$$(Ax, y)_\ell = (x, A^t y)_\ell \quad \forall x, y \in \mathbb{R}^n.$$

Definition 4. Given $A \in L(\mathbb{V})$, $B \in L(\mathbb{V})$ is called the adjoint of A with respect to the inner product (\cdot, \cdot) , if

$$(Au, v) = (u, Bv).$$

We denote $B = A'$. A is said to be self-adjoint or symmetric if $A' = A$. A is said to be SPD (symmetric positive definite) if A is symmetric and $(Av, v) > 0$ for any $v \in \mathbb{V} \setminus \{0\}$.

Lemma 2. Assume that $A \in L(\mathbb{V})$ is symmetric with respect to (\cdot, \cdot) . Then

1. $\sigma(A) \subset \mathbb{R}$.
2. If $Av_i = \lambda_i v_i$, $\lambda_1 \neq \lambda_2$, then $(v_1, v_2) = 0$.
3. A has a complete set of eigenvectors.

Exercise 1. Prove that A is self-adjoint if and only if A has all real eigenvalues and a complete set of eigenvectors (namely these eigenvectors form a basis of \mathbb{V}).

Definition 5 (A-inner product). Suppose $A \in \mathbb{R}^{n \times n}$ is an SPD matrix,

$$(1.10) \quad (x, y)_A = (Ax, y) = \sum_{i,j=1}^n a_{ij} x_i x_j$$

is defined as an A -inner product of x, y in \mathbb{R}^n .

Then A -norm and A -orthogonality can be defined successively.

Definition 6 (A-norm). Suppose $A \in \mathbb{R}^{n \times n}$ is an SPD matrix, for any $x \in \mathbb{R}^n$

$$\|x\|_A^2 = (x, x)_A = (Ax, x)$$

is defined as the A -norm of vector x .

Definition 7 (A-orthogonal). Suppose $A \in \mathbb{R}^{n \times n}$ is an SPD matrix, for any $x, y \in \mathbb{R}^n$, x is A -orthogonal to y which is denoted by $x \perp_A y$ is defined by

$$(x, y)_A = 0$$

or

$$x^T A y = 0.$$

Proposition 5. Assume $A \in L(\mathbb{V})$ is SPD w.r.t (\cdot, \cdot) . Then

1. Assume that “ \ast ” and “ \ast ” denote the transposes with respect to (\cdot, \cdot) and $(\cdot, \cdot)_A$ respectively, then

$$(BA)^\ast = B^t A^t.$$

2. BA is symmetric with respect to $(\cdot, \cdot)_A$ iff B is symmetric with respect to (\cdot, \cdot) .
3. If B is SPD with respect to the same inner product (\cdot, \cdot) , BA is SPD self-adjoint with respect to either $(A\cdot, \cdot)$ or $(B^{-1}\cdot, \cdot)$.

Proposition 6. Assume $A \in L(\mathbb{V})$ is SPD w.r.t (\cdot, \cdot) . Then $\sigma(A) \subset (0, \infty)$, $(\cdot, \cdot)_A = (A\cdot, \cdot)$ defines another inner product on \mathbb{V} and

$$(1.11) \quad \lambda_{\min}(A) = \min_{v \in \mathbb{V}} \frac{(Av, v)}{\|v\|^2} \quad \text{and} \quad \lambda_{\max}(A) = \max_{v \in \mathbb{V}} \frac{(Av, v)}{\|v\|^2}.$$

Proposition 7. Assume that $A : \mathbb{V} \mapsto \mathbb{V}$ and $\|\cdot\|$ is induced by an inner product (\cdot, \cdot) on \mathbb{V} . If A is symmetric with respect to (\cdot, \cdot) ,

$$\|A\| = \rho(A).$$

In general,

$$\|A\| = \rho(A'A)^{\frac{1}{2}}.$$

Here A' is the adjoint operator of A with respect to (\cdot, \cdot) .

Example 8. Given $\gamma \in (0, 1)$, consider the symmetric matrix

$$A = (\gamma^{|i-j|}) \in \mathbb{R}^{n \times n},$$

then

$$\rho(A) \leq \|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n \gamma^{|i-j|} \leq \frac{2}{1-\gamma}.$$

Thus

$$\sum_{i,j=1}^n \gamma^{|i-j|} x_i x_j \leq \frac{2}{1-\gamma} \sum_{i=1}^n x_i^2.$$

By taking $n \rightarrow \infty$, we have

$$\sum_{i,j=1}^{\infty} \gamma^{|i-j|} x_i x_j \leq \frac{2}{1-\gamma} \sum_{i=1}^{\infty} x_i^2.$$

More generally

$$(1.12) \quad \sum_{i,j=1}^{\infty} \gamma^{|i-j|} x_i y_j \leq \frac{2}{1-\gamma} \left(\sum_{i=1}^{\infty} x_i^2 \right)^{1/2} \left(\sum_{j=1}^{\infty} y_j^2 \right)^{1/2}.$$

Exercise 2. If $A \in L(\mathbb{V})$ is self-adjoint with respect to (\cdot, \cdot) and it is semi-definite, i.e. $(Av, v) \geq 0 \quad \forall v \in \mathbb{V}$. Then

$$(Au, v) \leq (Au, u)^{\frac{1}{2}} (Av, v)^{\frac{1}{2}} \quad \forall u, v \in \mathbb{V}.$$

Exercise 3. Under the assumption of Exercise 2, prove that, for any $\alpha \in [0, 1]$

$$\|A^{\frac{1+\alpha}{2}} v\| \leq (Av, v)^{\frac{1-\alpha}{2}} \|Av\|^{\alpha} \quad \forall v \in \mathbb{V}.$$

The following simple result is useful in multigrid analysis.

Exercise 4. Assume that $K : \mathbb{V} \mapsto \mathbb{V}$ is a symmetric operator with $\sigma(K) \subset [0, 1]$, then

$$((I - K)K^m v, K^m v) \leq \frac{1}{2m} (\|v\|^2 - \|K^m v\|^2).$$

1.5 Spectrum of a special tri-diagonal matrix

We shall now give an example on the spectrum of symmetric tri-diagonal matrix of the following form:

$$A = \text{diag}(b, a, b)$$

namely $a_{ii} = a$ for $1 \leq i \leq N$ and $a_{i,i-1} = a_{i-1,i} = b$ for $2 \leq i \leq N$, and all other entries of A are zero.

Proposition 8. *The eigenvalues of the symmetric tri-diagonal matrix $A = \text{diag}(b, a, b)$, for any two real numbers a, b , are given by*

$$\lambda_k = a + 2b \cos \frac{k\pi}{N+1}, \quad 1 \leq k \leq N$$

and the eigenvectors $w_k, 1 \leq k \leq N$, are given by

$$w_{k,j} = \sin \frac{jk\pi}{N+1}, \quad 1 \leq j \leq N$$

where $w_{k,j}$ is the j th component of the k th vector w_k .

Proof. Assume $x \in \mathbb{R}^N$ is an eigenvector to an eigenvalue $\lambda: Ax = \lambda x$. Then

$$(1.13) \quad bx_{j-1} + ax_j + bx_{j+1} = \lambda x_j, \quad 1 \leq j \leq N$$

$$(1.14) \quad x_0 = x_{N+1} = 0.$$

The characteristic equation of the above finite difference equation is

$$b\mu^2 + (a - \lambda)\mu + b = 0.$$

It is easy to see any eigenvalue of A must satisfy $|\lambda - a| \leq 2|b|$, the roots of the above equation are then given by

$$\mu = \frac{\lambda - a}{2b} \pm i \sqrt{1 - \left(\frac{\lambda - a}{2b}\right)^2} = e^{\pm i\theta}$$

where $i = \sqrt{-1}$ and θ is such that

$$\cos \theta = \frac{\lambda - a}{2b}, \quad \& \quad \text{namely } \lambda = a + 2b \cos \theta.$$

A solution of the difference equation (1.13) is given by

$$x_j = \text{Im}[e^{i\theta}]^j = \sin j\theta.$$

In order that (1.14) to be satisfied, θ should be chosen so that

$$\sin(N+1)\theta = 0$$

which gives

$$\theta = \frac{k\pi}{N+1}, \quad 1 \leq k \leq N.$$

The desired result then follows easily. \square

1.6 Matrix tensor products

In this section, we shall give a brief introduction of matrix tensor products which will be useful in the discussion of finite element equations on certain regular domains in two and three dimensions.

Definition 8. Given $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{p \times q}$, the tensor product of A and B , denoted by $A \otimes B$, is an $mp \times nq$ matrix or a $m \times n$ block matrix defined by

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \cdots & \cdots & \cdots & \cdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix}$$

By definition we have

- Theorem 2.**
1. $0 \otimes A = A \otimes 0 = 0$.
 2. $(A_1 + A_2) \otimes B = A_1 \otimes B + A_2 \otimes B$, $A \otimes (B_1 + B_2) = A \otimes B_1 + A \otimes B_2$.
 3. $(\alpha A) \otimes (\beta B) = (\alpha\beta)A \otimes B \quad \forall \alpha, \beta \in \mathbb{R}$.
 4. $(A \otimes B) \otimes C = A \otimes (B \otimes C)$.

The following less obvious result is important.

- Theorem 3.**
1. $(A_1 \otimes B_1)(A_2 \otimes B_2) = (A_1 A_2) \otimes (B_1 B_2)$.
 2. $(A \otimes B)^{-1} = A^{-1} \otimes B^{-1}$ if both A and B are invertible.
 3. $(A \otimes B)^t = A^t \otimes B^t$.
 4. The product of two lower (upper) triangular matrices is still lower (upper) triangular.
 5. The product of two orthogonal matrices is still orthogonal.

Proof. The first identity again follows from the definition and the second identity follows from the first one. The proof for other identities is also straightforward. \square

Theorem 4. Let $\phi(s, t) = \sum_{i,j=0}^p \alpha_{ij} s^i t^j$, $p \geq 0$. Define

$$\phi(A, B) = \sum_{i,j=0}^p \alpha_{ij} A^i \otimes B^j.$$

If $\{\lambda_i : i = 1 : m\}$ and $\{\mu_j : j = 1 : n\}$ are the spectral sets of A and B respectively, then the spectral set of $\phi(A, B)$ is $\{\phi(\lambda_i, \mu_j) : i = 1 : m, j = 1 : n\}$. Furthermore if $Ax = \lambda x$ and $By = \mu y$, then $\phi(A, B)(x \otimes y) = \phi(\lambda, \mu)(x \otimes y)$.

Proof. Assume that P and Q are the orthogonal matrices such that

$$P^* A P = S, \quad Q^* B Q = T$$

where S and T are both upper triangular matrices. Note that $S \otimes T$ is still upper triangular and $P \otimes Q$ is still orthogonal, and

$$(P \otimes Q)^*(A^i \otimes B^j)(P \otimes Q) = (P^* A^i P) \otimes (Q^* B^j Q) = S^i \otimes T^j.$$

This identity implies that $(P \otimes Q)^* \phi(A, B)(P \otimes Q) = \phi(S, T)$. Since $\phi(S, T)$ is an upper triangular matrix whose diagonal entries consist of $\phi(\lambda_i, \mu_j)$, this proves that $\{\phi(\lambda_i, \mu_j) : i = 1 : m, j = 1 : n\}$ is the spectral set of $\phi(A, B)$. The rest of the proof is straightforward by definition. \square

Example 9. Consider the block triadiagonal matrix

$$A = \text{diag}(-I, B, -I) \quad \& \quad \text{with } B = \text{diag}(-1, 4, -1).$$

Obviously

$$A = I \otimes B + C \otimes I \quad \text{with} \quad C = \text{diag}(-1, 0, -1).$$

By Proposition 8, the eigenvalues of B and C are as follows:

$$\lambda_i(B) = 4 - 2 \cos \frac{i\pi}{2(N+1)}, \quad \lambda_j(C) = -2 \cos \frac{j\pi}{2(N+1)}, \quad 1 \leq i, j \leq N.$$

By Theorem 4, the eigenvalues of A are given by

$$\begin{aligned} \lambda_{ij}(A) &= 4 - 2 \cos \frac{i\pi}{N+1} - 2 \cos \frac{j\pi}{N+1} \\ &= 4 \left(\sin^2 \frac{i\pi}{2(N+1)} + \sin^2 \frac{j\pi}{2(N+1)} \right). \end{aligned}$$

Example 10. Consider the block triadiagonal matrix which is the stiffness matrix for the center finite difference scheme for $-\epsilon u_{xx} - u_{yy} = f$

$$A = \text{diag}(-I, B, -I) \quad \& \quad \text{with } B = \text{diag}(-\epsilon, 2(1+\epsilon), -\epsilon).$$

Obviously

$$A = I \otimes B + C \otimes I \quad \text{with} \quad C = \text{diag}(-1, 0, -1).$$

By Proposition 8, the eigenvalues of B and C are as follows:

$$\lambda_i(B) = 2(1+\epsilon) - \epsilon \cos \frac{i\pi}{(N+1)}, \quad \lambda_j(C) = -2 \cos \frac{j\pi}{(N+1)}, \quad 1 \leq i, j \leq N.$$

By Theorem 4, the eigenvalues of A are given by

$$\begin{aligned} \lambda_{ij}(A) &= 2(1+\epsilon) - 2\epsilon \cos \frac{i\pi}{(N+1)} - 2 \cos \frac{j\pi}{N+1} \\ &= 2\epsilon \left(1 - \cos \frac{i\pi}{(N+1)} \right) + 2 \left(1 - \cos \frac{j\pi}{N+1} \right) \\ &= 4\epsilon \sin^2 \frac{i\pi}{2(N+1)} + 4 \sin^2 \frac{j\pi}{2(N+1)}. \end{aligned}$$

with eigenvectors

$$\phi_{ij}^{k\ell} = \sin \frac{k i \pi}{(N+1)} \sin \frac{\ell j \pi}{(N+1)}$$

1.7 Dual space

Given a vector space V , its algebraic dual space consists of all linear functionals on V together with a naturally induced linear structure.

In general, the notation $\langle \cdot, \cdot \rangle$ is used to denote a dual operation between V and V' . For example, $\langle f, v \rangle$, for $v \in V$ and $f \in V'$. There are a lot of ambiguities in the use of such a notation. In this book, to avoid unnecessary confusion, we only use such a notation in two major different ways.

1. Finite dimensional case, $\dim V < \infty$. If V is a finite dimensional space, we will only consider two different cases

(1) $V = \mathbb{R}^n$. In this case, we always identify $V' = \mathbb{R}^n$, and we always use

$$\langle f, v \rangle = (f, v)_E.$$

(2) $V \subset L^2(\Omega)$. In this case, V is finite dimensional functional space on a domain $\Omega \subset \mathbb{R}^d$. In this case, we always identify $V' = V$ and $\langle f, v \rangle = \int_{\Omega} f v dx$.

2. Infinitely dimensional case, $\dim V = \infty$. In this case, the situation is a little bit subtle. We consider two different cases, namely

(1) $V = \{(a_1, a_2, \dots, a_n, \dots)^T\}$ is a space of infinitely sequences. In this case, we always take $V' = \{(b_1, b_2, \dots, b_n, \dots)^T\}$ a space of infinitely sequences, and we always assume

$$\langle f, v \rangle = \sum_{i=1}^{\infty} f_i v_i = (f, v)_E$$

if $f = (f_1, f_2, \dots, f_n, \dots) \in l^2 \cap V'$, $v = (v_1, v_2, \dots, v_n, \dots) \in l^2 \cap V$.

Example 11. $(l^p)' = l^q$, $\frac{1}{p} + \frac{1}{q} = 1$.

(2) V is a space of function (or generalized function) on a domain $\Omega \subset \mathbb{R}^d$, we always assume

$$\langle f, v \rangle = \int_{\Omega} f v dx$$

if $v \in V \cap L^2(\Omega)$, $f \in V' \cap L^2(\Omega)$.

Example 12. $L^p(\Omega)' = L^q(\Omega)$, $\frac{1}{p} + \frac{1}{q} = 1$.

Example 13. $V = H_0^1(\Omega)$, $V' = H^{-1}(\Omega)$.

Example 14. $V = \mathcal{D}(\Omega)$, $V' = \mathcal{D}'(\Omega)$.

1.7.1 Finite dimensional spaces and dual basis

Let V be a finite dimensional space with

$$n = \dim V < \infty$$

and a basis $\{\phi_i\} \subset V$. It is easy to see that the dual of a finite dimensional space is also finite dimensional and

$$(1.15) \quad \dim V' = \dim V.$$

In fact, for each i , we can define a functional $\phi'_i \in V'$ such that:

$$\langle \phi'_j, \phi_i \rangle = \delta_{ij}, \quad \forall 1 \leq i, j \leq n.$$

Namely, if $v = \sum_{j=1}^n \alpha_j \phi_j$,

$$\langle \phi'_i, v \rangle = \alpha_i$$

and, for any $f \in V'$, we have,

$$\langle f, v \rangle = \sum_{i=1}^m \alpha_i \langle f, \phi_i \rangle = \sum_{i=1}^m \langle f, \phi_i \rangle \langle \phi'_i, v \rangle = \langle \sum_{i=1}^m \langle f, \phi_i \rangle \phi'_i, v \rangle$$

As a result

$$f = \sum_{i=1}^m \langle f, \phi_i \rangle \phi'_i.$$

This means (ϕ'_i) is a basis for V' . Hence (1.15) is valid. In fact, we identify $V' = V$ and hence we can view that both ϕ'_j and ϕ_j belong to V and are two basis of V , but they play different roles.

Lemma 3. Assume that (ϕ_i) and (ϕ'_i) are dual bases of V and V' respectively. Then

$$(1.16) \quad v = \sum_{i=1}^n \langle \phi'_i, v \rangle \phi_i, \quad v \in V,$$

and

$$(1.17) \quad f = \sum_{i=1}^n \langle f, \phi_i \rangle \phi'_i.$$

1.7.2 Vector and matrix representations

Given any $v \in V$, it can be uniquely written as a linear combination of the basis of V :

$$v = \sum_{i=1}^n v_i \phi_i.$$

The coefficient vector of this representation $\tilde{v} = (v_1, \dots, v_n)^T \in \mathbb{R}^n$ will be called the *vector representation* of v under the basis (ϕ_i) . This defines the following operator:

$$\tilde{\cdot}: V \mapsto \mathbb{R}^n.$$

which gives the following adjoint operator

$$\sim : (\mathbb{R}^n)' \mapsto V'$$

such that

$$\tilde{y}' = \sum_{i=1}^n y_i \phi'_i.$$

Namely

$$(y, \tilde{v})_{\ell^2} = \langle \tilde{y}', v \rangle.$$

Given two linear vector spaces V and W and $A \in \mathcal{L}(V, W)$, the *matrix representation* of A with respect to a basis (ϕ_1, \dots, ϕ_m) of V and a basis (ψ_1, \dots, ψ_n) of W is the matrix $\tilde{A} = (\tilde{a}_{ij}) \in \mathbb{R}^{n \times m}$ satisfying

$$(1.18) \quad (A\phi_1 \cdots, A\phi_m) = (\psi_1 \cdots, \psi_n) \tilde{A}.$$

Namely

$$(1.19) \quad A\phi_j = \sum_{i=1}^n \tilde{a}_{ij} \psi_i \quad 1 \leq j \leq m.$$

Thus

$$(1.20) \quad \tilde{a}_{ij} = \langle \psi'_i, A\phi_j \rangle$$

In the special case that $W = V'$, we always take $(\psi_1, \dots, \psi_n) = (\phi'_1, \dots, \phi'_n)$, the basis of V' that is dual to (ϕ_1, \dots, ϕ_n) , hence we have $\tilde{a}_{ij} = \langle \phi_i, A\phi_j \rangle$.

Obviously different bases of V or W correspond to different matrix representations of $A \in \mathcal{L}(V, W)$ and different vector representation of $v \in V$. But in another way the matrix representation can be viewed uniquely for a fixed base (ϕ) of the linear vector space V . Furthermore, any $v \in V$ can be represented explicitly with the help of basis of its dual space. In terms of the basis (ϕ'_i) that is dual to (ϕ_i) , by (1.16), we have

$$v_i = \langle \phi'_i, v \rangle.$$

Let us summarize some important properties of matrix representation and vector representation in the following lemma.

Lemma 4. *Let $A : V \mapsto V'$, then*

$$\langle Au, v \rangle = (\tilde{A}\tilde{u}, \tilde{v})_{\ell^2}.$$

Proof. It suffice to prove that for each i we have

$$\langle Au, \phi_i \rangle = (\tilde{A}\tilde{u})_i.$$

In fact, let $u = \sum_{j=1}^m \tilde{u}_j \phi_j$, then

$$\begin{aligned}\langle Au, \phi_i \rangle &= \langle A \sum_{j=1}^n \tilde{u}_j \phi_j, \phi_i \rangle = \sum_{j=1}^n \tilde{u}_j \langle A \phi_j, \phi_i \rangle \\ &= \sum_{j=1}^n \tilde{a}_{ij} \tilde{u}_j = (\widetilde{Au})_i.\end{aligned}$$

□

Lemma 5. Assume that U, V and W are linear vector spaces. The matrix representation and vector representation satisfy the following properties:

1. For $A \in \mathcal{L}(V, V)$,

$$\sigma(\widetilde{A}) = \sigma(A).$$

2. If $A \in \mathcal{L}(V, W)$ is invertible, then

$$\widetilde{A^{-1}} = \widetilde{A}^{-1}.$$

3. For $A \in \mathcal{L}(V, W)$ and $v \in V$,

$$\widetilde{Av} = \widetilde{A}\widetilde{v}.$$

4. For $A \in \mathcal{L}(V, W)$,

$$\widetilde{A^t} = (\widetilde{A})^t.$$

5. For $A \in \mathcal{L}(V, W)$ and $B \in \mathcal{L}(W, U)$,

$$\widetilde{BA} = \widetilde{B}\widetilde{A}.$$

1.7.3 Duals of operators

Given any vector space V , we can define the following mapping:

$$J : V \mapsto V''$$

by

$$\langle J(v), f \rangle = \langle f, v \rangle, \quad v \in V, f \in V'.$$

It is easy to see that J is injective. When J is also surjective, we then call V is reflexive. In this case, by our definition, V is simply a representation of V'' with the natural pairing

$$\langle v, f \rangle = \langle f, v \rangle, \quad v \in V, f \in V'.$$

In this case, V is called reflexive, and we simply write

$$V'' = V.$$

It is easy to see that every Hilbert space is reflexive. Examples of non-Hilbert spaces that are reflexive include ℓ^p and $L^p(\Omega)$ for $1 < p < \infty$. One well-known non-reflexive Banach space is $L^\infty(\Omega)$.

Unless otherwise specified, only reflexive spaces will be discussed in relationship with duals in the rest of this book. In fact, we will mainly focus on Hilbert spaces.

Given two vector spaces V and W and a linear mapping

$$T : V \mapsto W,$$

the adjoint of T :

$$T' : W' \mapsto V'$$

is defined by

$$\langle T'g, v \rangle = \langle g, Tv \rangle, \forall g \in W', v \in V.$$

In particular, we will be interested in linear mappings between V and its dual V' . Since we are mainly considering Hilbert spaces or finite dimensional spaces, we will thus assume that

$$V'' = V.$$

As a consequence, for any linear mappings $A : V \mapsto V'$ and $B : V' \mapsto V$, their duals $A' : V \mapsto V'$ and $B' : V' \mapsto V$ satisfying

$$(1.21) \quad \langle A'u, v \rangle = \langle Av, u \rangle, \quad \langle f, B'g \rangle = \langle g, Bf \rangle, \quad \forall u, v \in V, f, g \in V'.$$

1.8 Symmetric positive definite operators

We say that A (or B) is symmetric if $\langle A'u, v \rangle = \langle Av, u \rangle$, $\forall u, v \in V$ (or $\langle f, B'g \rangle = \langle g, Bf \rangle$, $\forall f, g \in V'$). We say that A is symmetric positive definite if

$$(u, v)_A \equiv \langle Au, v \rangle$$

defines an inner product on V which induces a norm $\|\cdot\|_A$ given by

$$\|v\|_A^2 = (v, v)_A.$$

We will call this inner product as energy inner product or A-inner product. Since we will use this inner product frequently in this book, we introduce an abbreviated notation for this inner product:

$$(1.22) \quad (u, v)_A \equiv \langle Au, v \rangle$$

We sometimes also use the notation $a(\cdot, \cdot)$ to denote the A-inner product. Thus, we have the following notation to denote the same thing:

$$a(u, v) = (u, v)_A = \langle Au, v \rangle$$

We will use $*$ to denote adjoint operator with respect to this inner product, namely for $C : V \rightarrow V$, $C^* : V \rightarrow V$ is the unique operator satisfying

$$(1.23) \quad (C^*u, v)_A = (u, Cv)_A, \quad u, v \in V.$$

By definition, we have

Lemma 6. *If $A : V \rightarrow V'$ is SPD, then for any $B : V' \rightarrow V$*

$$(BA)^* = B'A'.$$

Namely

$$(BAu, v)_A = (u, B'Av)_A.$$

We note that we have introduced two different adjoint operations here:

1. the adjoint “ \prime ” is defined between V and its dual V' using the pairing between V and V' , and,
2. the adjoint “ $*$ ” is defined between V and itself using the energy inner product on V .

For the rest of this chapter, we will introduce the matrix representation of some important operator step by step.

Lemma 7. *For a given operator $A \in \mathcal{L}(V, V')$, its matrix representation (a_{ij}) in terms of dual bases*

$$A\phi_j = \sum_{i=1}^n a_{ij}\phi'_i$$

is just the so-called stiffness matrix of the operator A :

$$\tilde{A} = (a_{ij}) \in \mathbb{R}^{n \times n} \text{ with } a_{ij} = \langle A\phi_j, \phi_i \rangle := a(\phi_j, \phi_i).$$

1.9 Symmetric semi-positive definite operators

We say that A is symmetric semi-positive definite (SSPD) if

$$(u, v)_A \equiv \langle Au, v \rangle$$

nonnegative bilinear form on V which induces a semi-norm $\|\cdot\|_A$ given by

$$(1.24) \quad \|v\|_A = (v, v)_A^{1/2}.$$

Lemma 8. *The semi-norm $\|\cdot\|_A$ defined by (1.24) satisfies the following properties:*

1. $\|u\|_A = 0$ iff $u \in N(A)$.
2. $\|\cdot\|_A$ is a norm on the quotient space $V/N(A)$.

Proof. Let $u \in V$ be such that

$$(u, u)_A = 0.$$

Then we have

$$(u, u)_A = \min_{t \in \mathbb{R}} (u + tv, u + tv)_A, \quad \forall v \in V.$$

This immediately implies that

$$\langle Au, v \rangle = (u, v)_A = 0, \quad \forall v \in V.$$

Thus $Au = 0$, namely $u \in N(A)$. The remaining results of the lemma is straightforward. \square

1.10 Identifying operators between V and V'

When V is a finite dimensional space, its dual V' can often be identified with V through the following identifying operator:

$$\iota f = \sum_{i=1}^n \langle f, \phi_i \rangle \phi_i, \quad f \in V'.$$

Namely,

$$(1.25) \quad \iota : V' \mapsto V, \quad \iota \phi'_j = \phi_j \quad (1 \leq j \leq n).$$

By definitions,

$$\iota^{-1} v = \sum_{i=1}^n \langle \phi'_i, v \rangle \phi'_i, \quad v \in V.$$

and Note that

$$\tilde{v} = \tilde{\iota}' v'$$

From the definition we see that the matrix representation of ι is the identity matrix, namely $\tilde{\iota}' = I$.

Using the identifying operator ι , we can define an inner product on V as follows

$$(u, v)_{\iota^{-1}} = (\iota^{-1} u, v) = \langle \iota u, v \rangle.$$

By definition,

$$(u, v)_{\iota^{-1}} = (\tilde{u}, \tilde{v})_{\ell^2} = \sum_{i=1}^n \langle \phi'_i, u \rangle \langle \phi'_i, v \rangle.$$

Lemma 9. Assume that $A : V \mapsto V'$ is SPD. Then $\iota A : V \mapsto V$ is symmetric (and positive definite) with respect to both of the $(\cdot, \cdot)_A$ and $(\cdot, \cdot)_{\iota^{-1}}$ inner products.

Proof. By definition,

$$\iota A u = \sum_{i=1}^n \langle A u, \phi_i \rangle \phi_i$$

Let us try

$$(\iota A u, v)_A = \sum_{i=1}^n \langle A u, \phi_i \rangle \langle A v, \phi_i \rangle = (u, \iota A v)_A.$$

This means that $\iota A : V \mapsto V$ is a symmetric positive definite operator with respect to inner product $(\cdot, \cdot)_A$. Furthermore

$$(\iota A u, v)_{\iota^{-1}} = (u, v)_A = (u, \iota A v)_{\iota^{-1}}.$$

□

1.11 Subspaces

1.11.1 Subspaces

Given a vector space V and a closed subspace $V_0 \subset V$, we shall now discuss various relationships between V_0 , V and their duals. We assume that V_0 is equipped with a basis

$$\phi_1^0, \phi_2^0, \dots, \phi_{n_0}^0, \quad n_0 = \dim V_0.$$

1.11.2 Inclusion operator: $V_0 \mapsto V$ and prolongation matrix: $\mathbb{R}^{n_0} \mapsto \mathbb{R}^n$

Consider the inclusion operator another operator $i_0 \in \mathcal{L}(V_0, V)$ that

$$i_0 v_0 = v_0 \quad v_0 \in V_0.$$

Let $P = (p_{ij}) = (\langle \phi_i', \phi_j^0 \rangle) : \mathbb{R}^{n_0} \mapsto \mathbb{R}^n$ such that

$$\phi_j^0 = \sum_{i=1}^n p_{ij} \phi_i,$$

or

$$(\phi_1^0, \dots, \phi_{n_0}^0) = (\phi_1, \dots, \phi_n)P.$$

If

$$v_0 = \sum_i v_i^0 \phi_i^0 = \sum_i v_i \phi_i$$

Then

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = P \begin{pmatrix} v_1^0 \\ \vdots \\ v_{n_0}^0 \end{pmatrix},$$

Namely P is the matrix representation of the inclusion $i_0 : V_0 \mapsto V$.

The space V is endowed with the inner product $(\cdot, \cdot)_A$, and the Riesz operator can be defined. Here we give more structures between different linear vector spaces originated from the nested relation in hierarchical basis as follows.

V_0' versus V' : restriction

Given $f \in V'$, let $f_0 \in V_0'$ be such that

$$\langle f_0, v_0 \rangle = \langle f, i_0 v_0 \rangle$$

This define an operator: $i_0' : V' \mapsto V_0'$ with $f_0 = i_0' f$.

Indeed, as the notation implies, i_0' is the dual of i_0 . If we write

$$f_0 = \sum_{i=1}^{n_0} f_i^0 (\phi_i^0)', \quad f = \sum_{i=1}^n f_i \phi_i',$$

then we have

$$\begin{pmatrix} f_1^0 \\ \vdots \\ f_{n_0}^0 \end{pmatrix} = P^T \begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix}$$

$\mathcal{P}^T : \mathbb{R}^n \mapsto \mathbb{R}^{n_0}$ is the matrix representation of $i'_0 : V' \mapsto V'_0$.

1.11.3 General Galerkin-projection

We recall that $A : V \mapsto V'$ and $A_0 : V_0 \mapsto V'_0$ that satisfy

$$\langle Au, v \rangle = (u, v)_A, \quad \langle A_0 u_0, v_0 \rangle = (u_0, v_0)_A, \quad u, v \in V, u_0, v_0 \in V_0.$$

The Galerkin projection $P_0 : V \mapsto V_0$ is defined by

$$(P_0 u, v_0)_A = (u, v_0)_A, \quad v \in V, v_0 \in V_0.$$

We note that, by definition,

$$A_0 = i'_0 A i_0.$$

and

$$A_0 P_0 = i'_0 A.$$

As a result,

$$(1.26) \quad P_0 = (A_0)^{-1} i'_0 A.$$

This implies that

$$(1.27) \quad i'_0 A (I - P_0) = 0.$$

Taking the matrix representation, we get

$$\tilde{A}_0 \tilde{P}_0 = \tilde{i}'_0 \tilde{A} = P^T \tilde{A}.$$

Namely

$$\tilde{P}_0 = (\tilde{A}_0)^{-1} P^T \tilde{A} = (P^T \tilde{A} P)^{-1} P^T \tilde{A}.$$

We note that

$$\tilde{P}_0 P = I.$$

Namely $\tilde{P}_0 : \mathbb{R}^n \mapsto \mathbb{R}^{n_0}$ is a left inverse of the prolongation matrix P .

Subspaces and duals

Given $V_0 \subset V$, let us ask the following simple question:

What is the relationship between V' and V'_0 ?

Given $f \in V'$, since $V_0 \subset V$, we clearly have a well-defined $\langle f, v_0 \rangle$ for any $v_0 \in V_0$. It is therefore tempting to make the following conclusion:

$$(1.28) \quad V_0 \subset V \Rightarrow V' \subset V'_0.$$

But, strictly speaking, the conclusion (1.28) is totally false. In fact, V'_0 and V' can not be compared because a functional in V' which is defined on V and a function V'_0 which is defined on V_0 can not be compared as they are defined on different spaces. This is analogous to the fact that the function spaces $L^2(-1, 1)$ and $L^2(0, 1)$ can not be compared to each other although we have

$$(0, 1) \subset (-1, 1)$$

because the domain is one key component that defines a function.

To further convince that the statement (1.28) is false, we consider the following simple example:

$$\mathbb{P}_n(0, 1) \subset L^2(0, 1).$$

The dual $(L^2(0, 1))' = L^2(0, 1)$, which is an infinite dimensional space, can not be a subspace of $(\mathbb{P}_n(0, 1))'$ which a finite dimensional.

Define

$$(1.29) \quad V_0^{(0)} = \{f \in V' : \langle f, v_0 \rangle = 0, \quad v_0 \in V_0.\}$$

Theorem 5. *The following isomorphic relationship holds:*

$$V'_0 \approx V'/V_0^{(0)}.$$

Proof. Define the restriction linear operator $R : V' \mapsto V'_0$ such that

$$\langle Rf, v_0 \rangle = \langle f, v_0 \rangle, \quad f \in V'.$$

By definition

$$\text{null}(R) = V_0^{(0)}$$

Thus

$$\overline{\text{range}(R)} = V/\text{null}(R) = V/V_0^{(0)}$$

It is easy to see that R is surjective (which is trivial in finite dimensional case and for infinite dimensional space, Hahn-Banach theorem is needed to establish this), namely

$$\overline{\text{range}(R)} = V_0$$

The desired result then follows. \square

Left :

$$\widetilde{i_0 v_0} = P\tilde{v}_0 \quad (\text{namely } \tilde{i}_0 = P),$$

and

$$\widetilde{P_0 v} = \tilde{P}_0 \tilde{v}.$$

Right :

$$P^T \tilde{f} = \tilde{i}'_0 f \quad (f \in V').$$

Back :

$$A_0 = i'_0 A i_0,$$

and

$$A_0 P_0 = i'_0 A,$$

and

$$i_0 \iota_0 i'_0 = \iota.$$

Front:

$$\tilde{A}_0 = P \tilde{A} P^T,$$

and

$$\tilde{A}_0 \tilde{P}_0 = P^T \tilde{A},$$

and

$$P \tilde{i}_0 P^T = \tilde{\iota}.$$

The three commutative diagrams on the top row mean that

$$\widetilde{i_0 v_0} = P\tilde{v}_0 (\text{namely } \tilde{i}_0 = P), A_0 = i_0 A i'_0, \tilde{i}'_0 = P^T.$$

The other two commutative diagram mean that:

$$\tilde{A} v = \tilde{A} \tilde{v}, \quad \tilde{A}_0 = P^T \tilde{A} P.$$

$$A u = f \iff \tilde{A} \tilde{u} = \tilde{f}.$$

$$\|v\|_A = \|\tilde{v}\|_{\tilde{A}}$$

$$R_{\text{sym}} = R' + R - R' A R$$

$$\tilde{R}_{\text{sym}} = (\tilde{R})' + \tilde{R} - \tilde{R}' \tilde{A} \tilde{R}$$

For Jacobi,

$$R f = \sum_{i=1}^n (\phi_i, \phi_i)_A^{-1} \langle f, \phi_i \rangle \phi_i$$

1.11.5 Finite element spaces and duals

Let us now consider the finite element space V_h defined by (2.5) in §2 with nodal basis functions given by (2.6). Thanks to the property (2.7), it is obvious that a basis function V'_h which is dual to the nodal basis functions (2.6) is given by

$$(1.30) \quad V'_h = \{\delta_{x_i} : i = 1 : N\}$$

where δ_{x_i} is the Dirac δ -function associated with the nodal point x_i :

$$(1.31) \quad \langle \delta_{x_i}, v \rangle = v(x_i), \quad v \in V_h.$$

As we shall see later, the construction of a basis of a finite element V_h is actually determined by a basis function in V'_h which is often a priori given.

A so-called stiffness matrix for this operator with respect to the nodal basis function given by (2.6) is

$$A = (a(\phi_j, \phi_i)).$$

Global polynomial space

Now we consider the polynomial space V_n defined by (1.7) together with the nodal basis functions given by (1.8). Again, we obviously have

$$(1.32) \quad V'_h = \{\delta_{x_i} : i = 1 : n\}$$

where δ_{x_i} is the Dirac δ -function associated with the nodal point x_i :

$$(1.33) \quad \langle \delta_{x_i}, v \rangle = v(x_i), \quad v \in V_n.$$

At the first glance, we seem to have obtained the same dual spaces V'_h and V'_n as given in (1.30) and (1.32) respectively for two different spaces V_h and V_n . But the spaces given in (1.30) and (1.32) are two completely different spaces because that the δ -functions in (1.30) and (1.32) have different domains as given in (1.31) and (1.33) respectively.

1.11.6 Different inner products

In this section, we consider some special vector space V that is equipped with another inner product $[\cdot, \cdot]$.

In this case, we can define the Gram matrix as follows:

$$(1.34) \quad M = ([\phi_j, \phi_i])_{i,j}.$$

This matrix M is often called the mass matrix when $W = L^2(\Omega)$.

We can also define an operator $R : V \mapsto V'$ such that

$$\langle Ru, v \rangle = [u, v], \quad u, v \in V.$$

It is easy to see that the matrix representation of R is just the W-Gram matrix given by (1.34). Namely

$$\tilde{R} = M.$$

Now given a subspace $V_0 \subset V$, we can then define the corresponding M_0 and R_0 which satisfy

$$\tilde{R}_0 = M_0.$$

Special projection

Given $V_0 \subset V$, consider the following orthogonal projection $G_0 : V \mapsto V_0$ by

$$(1.35) \quad [G_0u, v_0] = [u, v_0], \quad u \in V, v_0 \in V_0.$$

Lemma 10. Define the imbedding $i'_0 : V_0 \rightarrow V$, then

$$G_0 = R_0^{-1}i'_0R.$$

Its matrix representation is given by

$$\tilde{G}_0 = M_0^{-1}P^T M.$$

Proof. The proof follows by definition:

$$[R_0^{-1}i'_0Ru, v_0] = \langle i'_0Ru, v_0 \rangle = \langle Ru, v_0 \rangle = [u, v_0].$$

□

Lemma 11. Let

$$X = A^{-1}R(I - G_0), V_0^\perp = (I - P_0)V.$$

Then $X : V_0^\perp \mapsto V_0^\perp$ is an isomorphism and furthermore

$$X^{-1} = (I - P_0)R^{-1}A.$$

Proof. By definition,

$$\langle RG_0u, v_0 \rangle = \langle Ru, v_0 \rangle, \quad u \in V, v_0 \in V_0.$$

Let $S = A^{-1}R$, then

$$(SG_0u, v_0)_A = (Su, v_0)_A, \quad u \in V, v_0 \in V_0$$

By definition

$$P_0S(I - G_0) = 0.$$

This means that

$$S(I - G_0) : V_0^\perp \rightarrow V_0^\perp.$$

Furthermore the following identity holds:

$$(S(I - G_0))^{-1} = (I - P_0)S^{-1}$$

□

There are three cases that are of special interests.

1. $[\cdot, \cdot]$ is the $L^2(\Omega)$ -inner product. In this case G_0 is just the L^2 projection, which is often denoted by $Q_0 : V \mapsto V_0$, and it satisfies

$$\tilde{Q}_0 = M_0^{-1} P^T M.$$

2. $[u, v] = (u, v)_A$. In this case, G_0 is just the Galerkin project $P_0 : V \mapsto V_0$ that satisfies

$$\tilde{P}_0 = \tilde{A}_0^{-1} P^T \tilde{A}$$

which coincide with (1.27).

3. Given an operator $S : V \mapsto V$ which is SPD with respect to A -inner product, we define

$$[u, v] = (Su, v)_A.$$

Then we have an isomorphism

$$S(I - G_0) : V_0^\perp \rightarrow V_0^\perp.$$

and

$$[S(I - G_0)]^{-1} = (I - P_0)S^{-1}.$$

This result is useful in analyzing the two-level AMG convergence.

1.12 Infinite dimensional space and topological dual

1.12.1 Some examples of infinite dimensional vector spaces

We consider the following infinite-dimensional vector space:

1. ℓ^p : the set of sequences $\mathbf{x} = (x_1, x_2, \dots)$ such that $\|\mathbf{x}\|_{\ell^p} < \infty$, where

$$\|\mathbf{x}\|_{\ell^p} := \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}.$$

2. $L^p(\Omega)$: the set of integrable functions with bounded L^p norm $\|\cdot\|_{L^p}$:

$$\|f\|_{L^p} := \left(\int_{\Omega} |f|^p dx \right)^{1/p}.$$

They are both normed Banach space.

1.13 Continuous dual space

When V is a topological vector space, its continuous dual space, denoted by V' , consists of all continuous linear functionals on V . In this book, we will only consider continuous dual space which will be simply referred as *dual space*.

The space V' is a topological vector space equipped with the weak $*$ topology, namely, in V' , $f_n \rightarrow f$ if and only if $\langle f_n, v \rangle \rightarrow \langle f, v \rangle$ for all $v \in V$.

When V is a Banach space equipped with a norm $\|\cdot\|_V$, a strong topology can be defined on V' through the following induced dual norm

$$\|f\|_{V'} = \sup_{v \in V} \frac{\langle f, v \rangle}{\|v\|_V}.$$

With this norm, V' is also a Banach space.

A strong topology can also be defined on a non-Banach, but complete locally convex space. But we will not use such kind of strong topology in this book.

The definition of V' is abstract and in general, we can not say what is like for $f \in V'$.

Example 15. $V = \mathbb{R}^n$. For this simple example, what is its dual space V' ?

Let us first take any vector $u, v \in \mathbb{R}^n$, and consider

$$(1.36) \quad \langle u, v \rangle_0 := \sum_{i=1}^n u_i v_i = (u, v)_{\ell^2}$$

which is obviously a continuous linear functional.

Example 16. $V = \mathbb{P}_{n-1}[x]$. Let us consider the space, $\mathbb{P}_{n-1}[x]$, of polynomials of degree less than or equal to n . Given $p \in \mathbb{P}_{n-1}[x]$

$$p(x) = \sum_{i=1}^n p_i x^{i-1}$$

and consider

$$(1.37) \quad \langle p, v \rangle_1 := \int_0^1 p(x) q_v(x) dx, \quad q_v(x) = \sum_{i=1}^n v_i x^{i-1},$$

which is also obviously a continuous linear functional.

By definition, we have

$$(1.38) \quad u \in V' \text{ and } p \in V'$$

Does this mean that

$$\mathbb{R}^n \subset V', \quad \mathbb{P}_{n-1}[x] \subset V'?$$

Before we address this question, let us do the following simple calculation:

$$\langle p, v \rangle_1 = \int_0^1 p(x)q_v(x)dx = \sum_{i,j=1}^n h_{ij}p_i v_j = (u_p, v)_{\ell^2} = \langle u_p, v \rangle_0.$$

where

$$u_p = Hp \text{ with } h_{ij} = \int_0^1 x^{i+j-2} dx = \frac{1}{i+j-1}.$$

For such an $u = u_p \in \mathbb{R}^n$, we have

$$u_p = p \text{ in } V'.$$

At the first glance, this identity is very strange as the left side is a vector in \mathbb{R}^n and right side is a polynomial in $\mathbb{P}_{n-1}[x]$. But as a linear functional on V , they are identical since

$$(1.39) \quad \langle u_p, v \rangle_1 = \langle p, v \rangle_0, \quad v \in V$$

where the left side is evaluated as in (1.36) and the right side is evaluated as in (1.37).

Indeed, we have

$$(1.40) \quad (\mathbb{R}^n)' = \{u \in \mathbb{R}^n : \langle u, v \rangle = (u, v)_{\ell^2}\}.$$

And we also have

$$(1.41) \quad (\mathbb{R}^n)' = \{p \in \mathbb{P}_{n-1}[x] : \langle p, v \rangle = (p, q_v)_{L^2(0,1)}\}.$$

By convention, we often drop the details on how the functional evaluation is done and simply write in short:

$$(1.42) \quad (\mathbb{R}^n)' = \mathbb{R}^n$$

or

$$(1.43) \quad (\mathbb{R}^n)' = \mathbb{P}_{n-1}[x]$$

Obviously it is a bit dangerous to write these short-hand identities as they would imply the following pathological identity:

$$\mathbb{R}^n = \mathbb{P}_{n-1}[x].$$

This identity would make more sense if we their respective functional evaluations are specified:

$$(\mathbb{R}^n)_{\langle \cdot, \cdot \rangle_1} = (\mathbb{P}_{n-1}[x])_{\langle \cdot, \cdot \rangle_0}.$$

1.13.1 Representation of dual space

Dual is an abstract concept. We can have a realistic realization or representation of this dual in terms of some specific definition of dual pairing.

Definition 9. We say that a topological vector space W is a representation of V' if the following three conditions are satisfied:

1. For any $w \in W$, there is a well-defined continuous linear functional $[w, v]$ for any $v \in V$.²
2. For any $f \in V'$, there is a unique $w \in W$ such that

$$[w, v] = \langle f, v \rangle, \quad v \in V.$$

3. The topology in W is equivalent to the weak-* topology of W induced by the pairing $[\cdot, \cdot]$.³

Obviously, for any V , representation of V' is not unique. For the space \mathbb{R}^n , we see that either \mathbb{R}^n or $\mathbb{P}_{n-1}[x]$ is a representation of $(\mathbb{R}^n)'$. These two different spaces are different in the way how the pairings are evaluated. We note that the pairing in (1.36) based on ℓ^2 inner product is much more natural/neutral than the one defined in (1.37). In fact, the pairing based on ℓ^2 inner-product can be deemed to be the most natural/neutral. It is conventional that the dual space, V' , is often represented by a functional space based on ℓ^2 or L^2 pairings. It is by this convention that we often write (1.42) but not (1.43).

Other conventional representation of dual spaces based on ℓ^2 or L^2 pairings are as follows:

$$(\ell^p)' = \ell^q, \quad (L^p(\Omega))' = L^q(\Omega), \quad p \in (1, \infty], q \in [1, \infty) \text{ such that } \frac{1}{p} + \frac{1}{q} = 1.$$

When $p = q = 2$, we have

$$(\ell^2)' = \ell^2, \quad (L^2(\Omega))' = L^2(\Omega).$$

In Sobolev spaces, L^2 is the most basic Hilbert space (that involves no derivatives). It can be viewed as the pivotal space between many nontrivial Sobolev spaces and their duals (representations based on L^2 -pairing). For examples, for $p \geq 2$ and integer $m \geq 1$

$$L^p(\Omega) \hookrightarrow L^2(\Omega) \hookrightarrow (L^p(\Omega))' = L^q(\Omega)$$

By (1.39), for the linear function p defined in (1.37), there is a $u_p \in \mathbb{R}^n$ such that

$$(u_p, v)_{\ell^2} = \langle p, v \rangle, \quad v \in \mathbb{R}^n.$$

This is a special case of the following general Riesz Representation Theorem.

Theorem 7. Assume that V is a Hilbert space with linear inner product $(\cdot, \cdot)_V$. Then for any $f \in V'$, there exists a unique $u \in V$ such that

$$(u, v)_V = \langle f, v \rangle, \quad v \in V$$

and, furthermore

$$\|u\|_V = \|f\|_{V'}.$$

² This linear functional can often be explicitly defined for w in some subspace of W .

³ In the case of Banach space, the original norm in W , $\|\cdot\|_W$, is equivalent to the dual norm:

$$\|w\|_W \approx \sup_{v \in V} \frac{[w, v]}{\|v\|_V}.$$

In our language, the Riesz representation simply says that one representation of the dual space of any Hilbert space V is simply V itself if the inner product of V is used as pairing. With the exception of ℓ^2 and $L^2(\Omega)$, the representation given by Riesz representation theorem is not the natural representation of a Hilbert space.

In most literature, the Riesz representation theorem is often interpreted as saying the dual space of a Hilbert space is isometrically isomorphic to itself.

A convention for the choice of natural pairing

Given a vector space V , we will adapt a convention that has been commonly used in the literature for the choice of pairing $\langle \cdot, \cdot \rangle$ for two kinds of vector spaces.

The first kind is discrete space, such as finite dimensional \mathbb{R}^n and infinitely dimensional ℓ^p , we choose the discrete ℓ^2 pairing:

$$(1.44) \quad \langle f, v \rangle = \sum_i f_i v_i$$

In this case, we often have the following relationship:

$$(1.45) \quad V \hookrightarrow \ell^2 \hookrightarrow V'.$$

or

$$(1.46) \quad V \leftrightarrow \ell^2 \leftrightarrow V'.$$

In the finite dimensional case,

$$(1.47) \quad V = V'.$$

In our study, we rarely use infinitely dimensional discrete space. Thus, unless otherwise specified, we always have the relationship (1.47).

Thus ℓ^2 is often known as the pivotal space between V and V' .

The second kind is infinitely dimensional functional space defined on a domain Ω , we choose the continuous L^2 pairing:

$$(1.48) \quad \langle f, v \rangle = \int_{\Omega} f(x)v(x)$$

In this case, we often have the following relationship (when Ω is bounded)

$$(1.49) \quad V \hookrightarrow L^2(\Omega) \hookrightarrow V'.$$

or

$$(1.50) \quad V \leftrightarrow L^2(\Omega) \leftrightarrow V'.$$

Thus $L^2(\Omega)$ is often known as the pivotal space between V and V' .

Unless specified otherwise, we will usually start with $V \subset L^2(\Omega)$, e.g. $V = H_0^1(\Omega)$, and thus (1.49) holds. Again, for finite dimensional functional space (such as finite element space), we always have

$$(1.51) \quad V = V'.$$

Dual basis when we view V as a representation of V'

We assume that V is equipped with an inner product $[\cdot, \cdot]$, then V is a representation of V' using the inner product $[\cdot, \cdot]$. In fact for any $f \in V'$, it is easy to see that there exists a unique $u \in V$ such that

$$[u, v] = \langle f, v \rangle, \quad v \in V.$$

Now, we call $\{\phi_j^*\} \subset V$ is a basis of V that is dual to the original space if

$$[\phi_j^*, \phi_i] = \delta_{ij}, \quad \forall 1 \leq i, j \leq n.$$

It is easy to see that such a dual basis in V uniquely exist.

The following identity holds:

$$v = \sum_{i=1}^n \langle \phi_i', v \rangle \phi_i = \sum_{i=1}^n [\phi_i^*, v] \phi_i, \quad \forall v \in V.$$