Geometry Processing (601.458/658)

Misha Kazhdan

Outline

Recall

Mathematical Preliminaries

Spectral Decomposition of Self-Adjoint Operators

Recall

Symmetric and Self-Adjoint Operators:

Given a vector space V, a linear map $L \in \mathrm{Hom}(V, V^*)$ is *symmetric* if:

$$[L(v)](w) = [L(w)](v) \qquad \forall v, w \in V$$
$$\Updownarrow$$
$$L = L^*$$

Given an inner-product space $\{V, B: V \to V^*\}$, an endomorphism $L \in \mathrm{End}(V)$ is *self-adjoint* if:

$$\langle v, L(w)\rangle_B = \langle L(v), w\rangle_B \qquad \forall v, w \in V$$
$$\Updownarrow$$
$$L^* \circ B = B \circ L$$

Recall

Symmetric and Self-Adjoint Operators:

For an inner-product space $\{V, B: V \to V^*\}$, an endomorphism $L \in \mathrm{End}(V)$ is self-adjoint if and only if $B \circ L$ is symmetric.

For an inner-product space $\{V, B: V \to V^*\}$, a linear map $L \in \mathrm{Hom}(V, V^*)$ is symmetric if and only if $B^{-1} \circ L$ is self-adjoint.

Recall

Given vector spaces V and W, and a linear map $L \in \mathrm{Hom}(V, W)$, the *kernel* of the linear map is the subspace of V consisting of vectors mapped to zero:

$$\mathrm{Ker}(L) = \{v \in V | L(v) = 0\}$$

Recall

Given a vector space V and a differentiable function $F: V \to \mathbb{R}$, the *differential* of F at $v \in V$, denoted $dF|_v \in V^*$, is an element of the dual space giving the local change of F at v:

$$dF\Big|_v(w) = \lim_{\varepsilon \to 0} \frac{F(v + \varepsilon \cdot w) - F(v)}{\varepsilon}$$

The function F is *regular*, a.k.a. *non-singular*, at $v \in V$ if its differential is non-zero at v.

The vector $v \in V$ is a local extremum of F only if the differential at v is zero.

Recall

Quadratic Forms:

Given a vector space V and a bilinear form $B \in \mathrm{Hom}(V, V^*)$, the associated *quadratic form* is the real-valued function on V:

$$Q_B : V \to \mathbb{R}$$
$$v \mapsto \frac{1}{2} B(v, v)$$

At $v \in V$, the differential of the quadratic form is:

$$dQ_B \Big|_V = B(v)$$

Recall

Inner-Product Subspaces:

Given an inner-product space $\{V, B: V \to V^*\}$ and a subspace $U \subset V$, we have the injection $\iota: U \hookrightarrow V$.

Using the injection, we can pull-back the inner-product to U:

$$B_U \equiv \iota^* \circ B \circ \iota$$

Because $B: V \to V^*$ is symmetric positive definite and $\iota: U \hookrightarrow V$ is injective, the bilinear form $B_U: U \to U^*$ is also symmetric positive definite.

$\Rightarrow$ The subspace U is an inner-product space, $\{U, B_U: U \to U^*\}$.

Recall

Inner-Product Subspaces:

Suppose $L \in \mathrm{End}(V)$ is self-adjoint and $U \subset V$ is a subspace *fixed* under L:

$$L: U \to U$$

Then for all $u, v \in U$ we have:

$$\begin{aligned} \langle v, L(w) \rangle_{B_U} &= \langle v, L(w) \rangle_B \\ &= \langle L(v), w \rangle_B \\ &= \langle L(v), w \rangle_{B_U} \end{aligned}$$

$\Rightarrow$ The restriction of the linear operator $L: U \to U$ is self-adjoint.

Recall

Linear Dependence:

Given a vector space V and vectors $\{v_1, \dots, v_m\}$, the vectors are *linearly dependent* if there exist $\alpha_1, \dots, \alpha_m \in \mathbb{R}$, not all zero, such that:

$$\alpha_1 \cdot v_1 + \cdots + \alpha_m \cdot v_m = 0$$

Note:

If $\{\alpha_1, \dots, \alpha_m\}$ express the linear dependence of the vectors $\{v_1, \dots, v_m\}$ then so do $\{\beta \cdot \alpha_1, \cdots, \beta \cdot \alpha_m\}$ for any non-zero $\beta \in \mathbb{R}$.

Outline

Recall

Mathematical Preliminaries

Spectral Decomposition of Self-Adjoint Operators

Mathematical Preliminaries

Claim:

Given a vector space V, if $l \in V^*$ is a dual vector, and $v \in V$ is such that $l(v) \neq 0$, any vector $w \in W$ is a multiple of v plus an element of the kernel:

$$w - \alpha_w \cdot v \in \mathrm{Ker}(l)$$

Proof:

Setting $\alpha_w = l(w)/l(v)$:

$$\begin{aligned} l(w - \alpha_w \cdot v) &= l(w) - \alpha_w \cdot l(v) \\ &= l(w) - \frac{l(w)}{l(v)} \cdot l(v) \\ &= l(w) - l(w) \\ &= 0 \end{aligned}$$

$\Rightarrow$ The difference between w and $\alpha_w \cdot v$ is in the kernel.

Mathematical Preliminaries

Claim:

Given two dual vectors $l_1, l_2 \in V^*$ with $l_1 \neq 0,$ the vectors are *linearly dependent* if and only if:

$$\mathrm{Ker}(l_1) \subset \mathrm{Ker}(l_2)$$

Proof:

Denote by $\hat{v} \in V$ some vector on which l_1 does not vanish:

$$l_1(\hat{v}) \neq 0$$

$\Rightarrow$ For any $w \in V$ there exists $u \in \mathrm{Ker}(l_1)$ such that:

$$w = \frac{l_1(w)}{l_1(\hat{v})} \cdot \hat{v} + u$$

For all $w \in V$, $\exists u \in \mathrm{Ker}(l_1)$ s.t.
$$w = \frac{l_1(w)}{l_1(\hat{v})} \cdot \hat{v} + u$$

Mathematical Preliminaries

Proof $(\Rightarrow)$: "If l_1 and l_2 are linearly dependent then…"

If the l_1 and l_2 are linearly dependent, there exist $\alpha_1, \alpha_2 \in \mathbb{R}$, not both zero:

$$\alpha_1 \cdot l_1 = \alpha_2 \cdot l_2$$

$\Rightarrow$ Applying to $\hat{v}$ gives:

$$\alpha_1 \cdot l_1(\hat{v}) = \alpha_2 \cdot l_2(\hat{v})$$

$\Rightarrow$ Since $l_1(\hat{v}) \neq 0$, we cannot have $\alpha_2 = 0$, and dividing gives:

$$\frac{\alpha_1}{\alpha_2} = \frac{l_2(\hat{v})}{l_1(\hat{v})}$$

$\Rightarrow$ We also have linear dependence with non-trivial coefficients:

$$l_2(\hat{v}) \cdot l_1 = l_1(\hat{v}) \cdot l_2$$

For all $w \in V$, $\exists u \in \mathrm{Ker}(l_1)$ s.t.
$$w = \frac{l_1(w)}{l_1(\hat{v})} \cdot \hat{v} + u$$

Mathematical Preliminaries

Proof $(\Rightarrow)$: "If l_1 and l_2 are linearly dependent then…"

$$l_2(\hat{v}) \cdot l_1 = l_1(\hat{v}) \cdot l_2$$

$\Rightarrow$ For all $u \in \mathrm{Ker}(l_1)$ we have:

$$0 = l_1(\hat{v}) \cdot l_2(u)$$

$\Rightarrow$ Since $l_1(\hat{v}) \neq 0$, for all $u \in \mathrm{Ker}(l_1)$ we have:

$$0 = l_2(u)$$

$\Rightarrow$ For all $u \in \mathrm{Ker}(l_1)$ we have:

$$u \in \mathrm{Ker}(l_2)$$

$\Rightarrow \mathrm{Ker}(l_1) \subset \mathrm{Ker}(l_2)$

For all $w \in V$, $\exists u \in \mathrm{Ker}(l_1)$ s.t.
$$w = \frac{l_1(w)}{l_1(\hat{v})} \cdot \hat{v} + u$$

Mathematical Preliminaries

Proof $(\Leftarrow)$: "If $\mathrm{Ker}(l_1) \subset \mathrm{Ker}(l_2)$ then..."

Consider the non-trivial linear combination:

$$l = l_2(\hat{v}) \cdot l_1 - l_1(\hat{v}) \cdot l_2$$

$\Rightarrow$ The evaluation at any $w \in V$ is:

$$\begin{aligned} l(w) &= l_2(\hat{v}) \cdot \left(\frac{l_1(w)}{l_1(\hat{v})} \cdot l_1(\hat{v}) + l_1(u) \right) - l_1(\hat{v}) \cdot \left(\frac{l_1(w)}{l_1(\hat{v})} \cdot l_2(\hat{v}) + l_2(u) \right) \\ &= l_2(\hat{v}) \cdot l_1(w) - l_1(w) \cdot l_2(\hat{v}) - l_1(\hat{v}) \cdot l_2(u) \\ &= -l_1(\hat{v}) \cdot l_2(u) \\ &= 0 \end{aligned}$$

since $\mathrm{Ker}(l_1) \subset \mathrm{Ker}(l_2)$.

$\Rightarrow$ The vectors l_1 and l_2 are linearly dependent.

For all $w \in V$, $\exists u \in \mathrm{Ker}(l_1)$ s.t.

$$w = \frac{l_1(w)}{l_1(\hat{v})} \cdot \hat{v} + u$$

Mathematical Preliminaries

Claim:

Given two dual vectors $l_1, l_2 \in V^*$ with $l_1 \neq 0,$ we have $\mathrm{Ker}(l_1) \subset \mathrm{Ker}(l_2)$ if and only if:

$$l_2(\hat{v}) \cdot l_1 = l_1(\hat{v}) \cdot l_2$$

Corollary:

Since $l_1(\hat{v}) \neq 0$, this means there exists $\lambda \in \mathbb{R}$ such that:

$$\lambda \cdot l_1 = l_2$$

Mathematical Preliminaries

Given a compact (closed and bounded) domain Ω, any *continuous* function $F: \Omega \to \mathbb{R}$ attains a maximum/minimum somewhere on Ω.

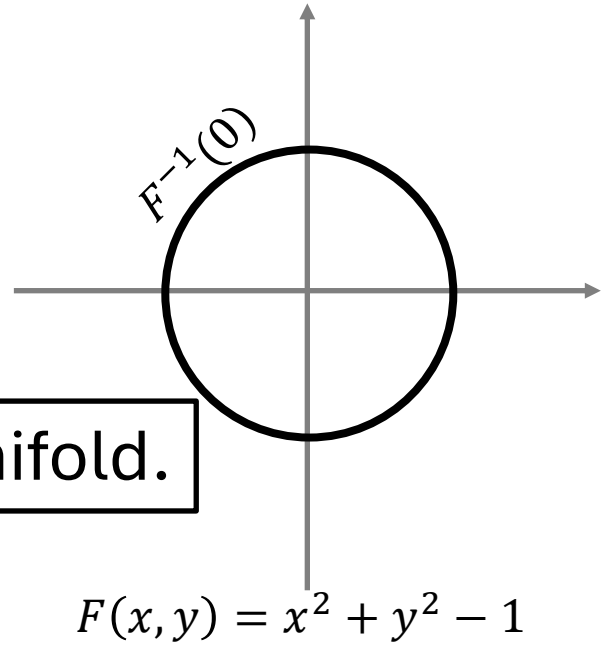
Mathematical Preliminaries

Level-Sets:

Given a vector space V and a differentiable (implicit) function $F : V \to \mathbb{R}$, the *zero level-set* of F is the set of vectors in V at which F vanishes:

$$F^{-1}(0) = \{v \in V | F(v) = 0\}$$

If the function is non-singular at all $v \in F^{-1}(0)$, the level-set is *manifold*.



In what follows, we assume that the level-set are manifold.

Mathematical Preliminaries

Definition:

Given an inner-product space $\{V, B: V \to V^*\}$ the *unit-sphere in* V is the zero level-set of the function:

$$F: V \to \mathbb{R}$$
$$v \mapsto \|v\|_B^2 - 1$$

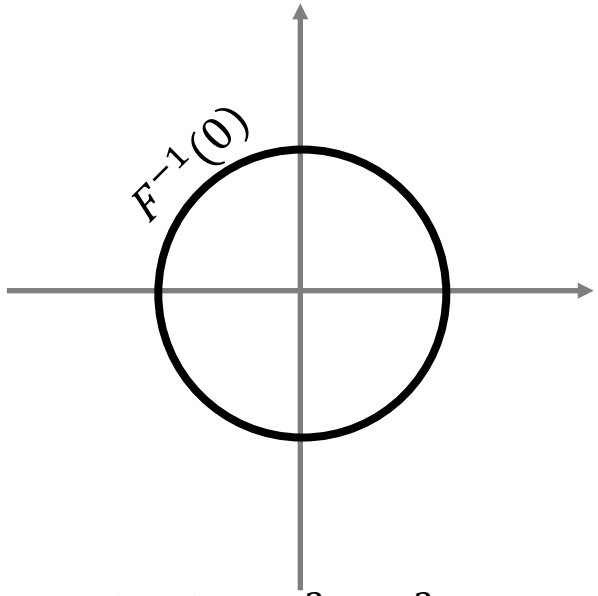
Noting that $F = 2 \cdot Q_B - 1$ and taking the differential:

$$dF\Big|_v = 2 \cdot B(v)$$

⇒ Since B is positive-definite, $B(v) \neq 0$ whenever $v \neq 0$.

⇒ The differential is non-singular for all v with $\|v\|_B^2 = 1$.

⇒ The unit sphere in V is manifold.

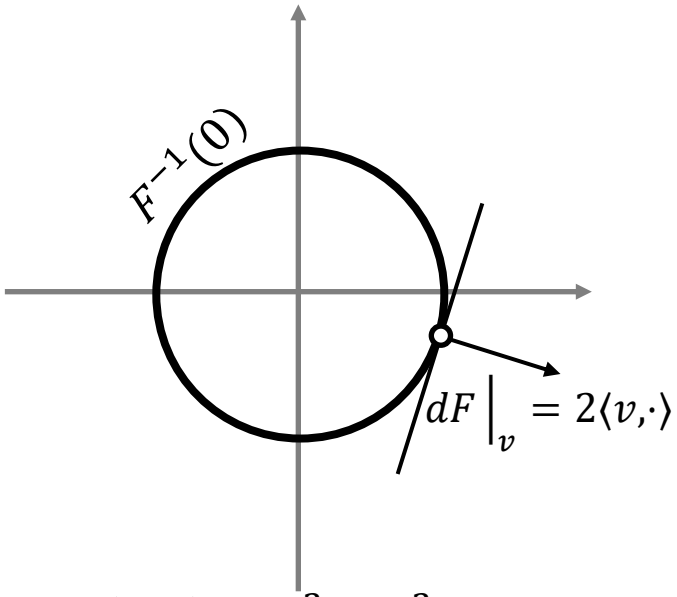


Mathematical Preliminaries

Implicit Surfaces:

Given an implicit function $F: V \to \mathbb{R}$, the *tangent space at* v is the subspace of V consisting of directions along which F does change (locally):

$$\begin{aligned} T_v F^{-1}(0) &= \left\{ w \in V \middle| \lim_{\varepsilon \to 0} \frac{F(v + \varepsilon \cdot w) - F(v)}{\varepsilon} = 0 \right\} \\ &= \{ w \in V | dF|_v(w) = 0 \} \\ &= \mathrm{Ker}\left(dF \Big|_v \right) \end{aligned}$$



Mathematical Preliminaries

Lagrange Multipliers:

Given an implicit function $F: V \to \mathbb{R}$, a function $G: V \to \mathbb{R}$ is extremized at $v \in F^{-1}(0)$ if the differentials at v are multiples of each other.

Proof Sketch:

G is extremized if it doesn't change along tangents:

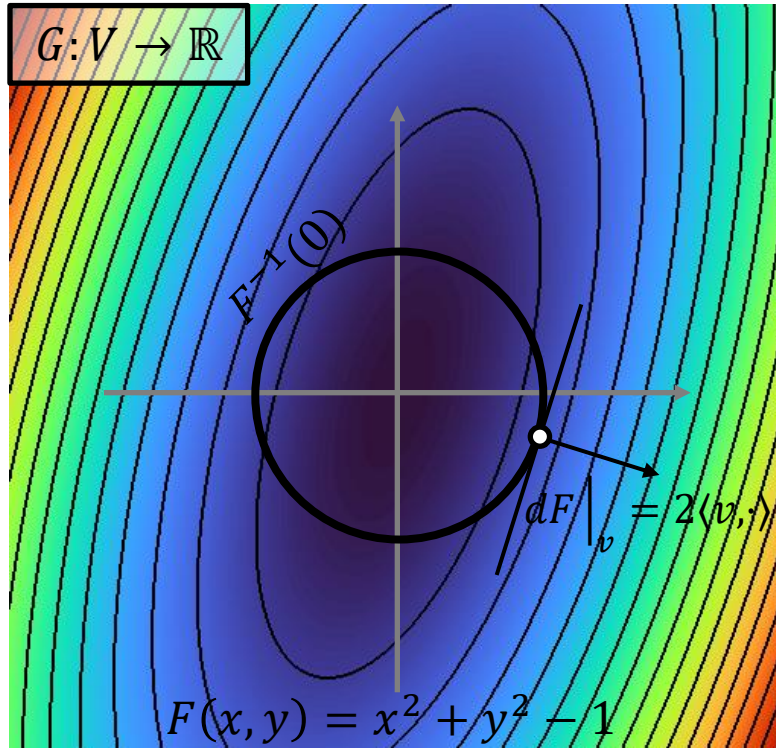
$$T_v F^{-1}(0) \subset \mathrm{Ker}\left(dG\Big|_v\right)$$

$$\Updownarrow$$

$$\mathrm{Ker}\left(dF\Big|_v\right) \subset \mathrm{Ker}\left(dG\Big|_v\right)$$

$$\Updownarrow$$

$$\lambda \cdot dF\Big|_v = dG\Big|_v$$



Outline

Recall

Mathematical Preliminaries

Spectral Decomposition of Self-Adjoint Operators

Eigenvalues

Definition:

Given a vector space V and an endomorphism $L \in \mathrm{End}(V)$, a vector $v \in V$ is an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if:

$$L(v) = \lambda \cdot v$$

Note that it does not make sense to talk about "eigenvalues" of linear maps between different vector spaces!

Eigenvalues

Definition:

Given a vector space V and an endomorphism $L \in \mathrm{End}(V)$, a vector $v \in V$ is an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if:

$$L(v) = \lambda \cdot v$$

Note:

If v_1 and v_2 are both eigenvectors with eigenvalue λ, for all $\alpha, \beta \in \mathbb{R}$:

$$\begin{aligned} L(\alpha \cdot v_1 + \beta \cdot v_2) &= \alpha \cdot L(v_1) + \beta \cdot L(v_2) \\ &= \alpha \cdot \lambda \cdot v_1 + \beta \cdot \lambda \cdot v_2 \\ &= \lambda \cdot (\alpha \cdot v_1 + \beta \cdot v_2) \end{aligned}$$

$\Rightarrow$ A linear combination of eigenvectors with the same eigenvalues is itself an eigenvector (with the same eigenvalue)

Eigenvalues

Definition:

Given a vector space V and an endomorphism $L \in \mathrm{End}(V)$, a vector $v \in V$ is an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if:

$$L(v) = \lambda \cdot v$$

Note:

If L is invertible and $v \in V$ is an eigenvector of L with eigenvalue $\lambda \in \mathbb{R}$, then λ cannot be zero. (Otherwise $L(v) = 0$ for $v \neq 0$, contradicting invertibility).

Eigenvalues

Definition:

Given a vector space V and an endomorphism $L \in \mathrm{End}(V)$, a vector $v \in V$ is an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if:

$$L(v) = \lambda \cdot v$$

Note:

If L is invertible and $v \in V$ is an eigenvector of L with eigenvalue $\lambda \in \mathbb{R}$:

$$\begin{aligned} v &= L^{-1}\big(L(v)\big) \\ &= L^{-1}(\lambda \cdot v) \\ &= \lambda \cdot L^{-1}(v) \end{aligned}$$

$$\Updownarrow$$

$$L^{-1}(v) = \frac{1}{\lambda} \cdot v$$

$\Rightarrow$ The vector v is an eigenvector of L^{-1}, with reciprocal eigenvalue.

Eigenvalues

Definition:

Given a vector space V and an endomorphism $L \in \mathrm{End}(V)$, a vector $v \in V$ is an *eigenvector* with *eigenvalue* $\lambda \in \mathbb{R}$ if:

$$L(v) = \lambda \cdot v$$

Note:

If $v \in V$ is an eigenvector of L with eigenvalue $\lambda \in \mathbb{R}$:

$$\begin{aligned}(L + \varepsilon \cdot \mathrm{Id}_V)(v) &= L(v) + \varepsilon \cdot \mathrm{Id}_V(v) \\ &= \lambda \cdot v + \varepsilon \cdot v \\ &= (\lambda + \varepsilon) \cdot v\end{aligned}$$

$\Rightarrow$ The vector v is also an eigenvector of $L + \varepsilon \cdot \mathrm{Id}_V$, with eigenvalue $\lambda + \varepsilon$.

Spectral Decomposition

Claim:

Given an inner-product space $\{V, B: V \to V^*\}$ and given a self-adjoint operator $L \in \mathrm{End}(V)$, there exists an **orthonormal** *eigen-basis* $\{v_1, \ldots, v_n\}$ (with associated eigenvalue $\{\lambda_1, \ldots, \lambda_n\}$):

$$L(v_i) = \lambda_i \cdot v_i$$

with:

$$\langle v_i, v_j \rangle_B = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

Proof:

First, we show that there exists one eigenvector.

Then we induct on the orthogonal subspace

Spectral Decomposition

Proof (Existence of an Eigenvector):

Given a self-adjoint $L \in \mathrm{End}(V)$, we can define the symmetric bilinear form:

$$B \circ L : V \to V^*$$

This defines the quadratic form:

$$Q_{B\circ L} : V \to \mathbb{R}$$
$$v \mapsto \frac{1}{2}\big[B\big(L(v)\big)\big](v) = \frac{1}{2}\langle v, L(v)\rangle_B$$

The quadratic form has differential:

$$dQ_{B\circ L}\Big|_v = B\big(L(v)\big)$$

$$Q_{B\circ L}(v) = \tfrac{1}{2}\langle v, L(v)\rangle_B$$
$$dQ_{B\circ L}\Big|_v = B\big(L(v)\big)$$

Spectral Decomposition

Proof (Existence of an Eigenvector):

Consider extremizing the quadratic form over the unit-sphere in V – the zero level-set set of $F(v) = \|v\|_B^2 - 1$.

$\Rightarrow$ At an extremum $v \in F^{-1}(0)$, the differentials are multiples of each other:

$$dQ_{B\circ L}\Big|_v = \lambda \cdot dF\Big|_v$$
$$\Updownarrow$$
$$B\big(L(v)\big) = 2 \cdot \lambda \cdot B(v)$$
$$\Updownarrow$$
$$L(v) = 2 \cdot \lambda \cdot v$$

$\Rightarrow$ There is an eigenvector of L at every extremum of $Q_{B\circ L}$ on the unit sphere.

$\Rightarrow$ Since the unit sphere is compact, $Q_{B\circ L}$ must attain an extremum on it.

$\Rightarrow$ There exists $v \in V$ that is an eigenvector of L.

Spectral Decomposition

Proof (Orthogonality of Eigenspaces):

Given $v_1 \in V$ an eigenvector of L (with eigenvalue λ_1) we would like to show that L takes the space perpendicular to v_1 into itself.

Denote by $V_1^{\perp} \subset V$ the subspace of vectors perpendicular to v:

$$V_1^{\perp} = \{w \in V | \langle v_1, w \rangle_B = 0\}$$

By self-adjointness, for any vector $w \in V_1$ we have:

$$\begin{aligned} \langle v_1, L(w) \rangle_B &= \langle L(v_1), w \rangle_B \\ &= \langle \lambda_1 \cdot v_1, w \rangle_B \\ &= \lambda_1 \cdot \langle v_1, w \rangle_B \\ &= 0 \end{aligned}$$

$\Rightarrow$ For all $w \in V_1^{\perp}$, we have $L(w) \in V_1^{\perp}$.

Spectral Decomposition

Proof (Orthogonality of Eigenspaces):

The self-adjoint operator L takes the space perpendicular to the eigenvector v_1 into itself:

$$L: V_1^{\perp} \to V_1^{\perp}$$

$\Rightarrow$ The restriction of L to the subspace $V_1^{\perp}$ is also self-adjoint.

$\Rightarrow$ The endomorphism L has an eigenvector in $v_2 \in V_1^{\perp}$.

$\Rightarrow$ The endomorphism L takes the subspace of vectors perpendicular to **both** v_1 and v_2 into itself, and its restriction to that space is also self-adjoint.

$\Rightarrow$ The endomorphism L has an eigenvector v_3 in the space of vectors perpendicular to both v_1 and v_2.

$\vdots$

$\Rightarrow$ There exists an orthonormal *eigen-basis* $\{v_1, \dots, v_n\}$ for L.

Arnoldi Method

In Practice:

Given an inner-product space, $\{V, B: V \to V^*\}$ and a self-adjoint operator $L \in \mathrm{End}(V)$ how do we compute the eigenvector of L with largest absolute eigenvalue?

Given an inner-product space, $\{V, B: V \to V^*\}$, a self-adjoint operator $L \in \mathrm{End}(V)$, and eigenvectors of L with largest absolute eigenvalues, how do we compute the next eigenvector?

Arnoldi Method

Approach (first eigenvector):

Choose a random vector $r^0 \in V$.

With respect to the eigen-basis/values $\{(v_1, \lambda_1), \dots, (v_n, \lambda_n)\}$ we have:

$$r^0 = \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n$$

Applying the linear operator L to the random vector gives:

$$\begin{aligned} r^1 = L(r^0) &= L(\alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n) \\ &= \alpha_1 \cdot L(v_1) + \cdots + \alpha_n \cdot L(v_n) \\ &= \alpha_1 \cdot \lambda_1 \cdot v_1 + \cdots + \alpha_n \cdot \lambda_n \cdot v_n \end{aligned}$$

Arnoldi Method

Approach (first eigenvector):

Choose a random vector $r^0 \in V$.

With respect to the eigen-basis/values $\{(v_1, \lambda_1), \dots, (v_n, \lambda_n)\}$ we have:

$$r^0 = \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n$$

Applying the linear operator L to the random vector gives:

$$r^1 = L(r^0) = \alpha_1 \cdot \lambda_1 \cdot v_1 + \cdots + \alpha_n \cdot \lambda_n \cdot v_n$$

Applying the linear operator L to the random vector k times gives:

$$\begin{aligned} r^k &= L(r^{k-1}) \\ &= \alpha_1 \cdot \lambda_1^k \cdot v_1 + \cdots + \alpha_n \cdot \lambda_n^k \cdot v_n \end{aligned}$$

Arnoldi Method

Approach (first eigenvector):

Choose a random vector $r^0 \in V$.

$$r^k = L^k(r^0) = \alpha_1 \cdot \lambda_1^k \cdot v_1 + \cdots + \alpha_n \cdot \lambda_{n-1}^k \cdot v_{n-1} + \alpha_n \cdot \lambda_n^k \cdot v_n$$

$$= \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n} \right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n} \right)^k + v_n \right)$$

Assume the eigen-basis is sorted by eigenvalue, $|\lambda_i| < |\lambda_{i+1}|$, then:

$$\lim_{k \to \infty} (\lambda_i / \lambda_n)^k = 0 \quad \forall\, 1 \leq i < n$$

In the limit, we have:

$$v_n = \lim_{k \to \infty} \frac{r^k}{\|r^k\|_B}$$

Arnoldi Method

Approach (first eigenvector):

$$r^k = L^k(r^0) = \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n} \right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n} \right)^k + v_n \right)$$

$$v_n = \lim_{k \to \infty} \frac{r^k}{\|r^k\|_B}$$

Since v_n is a unit-norm eigenvector, the associated eigenvalue is obtained by taking the inner-product:

$$\begin{aligned} \langle v_n, L(v_n) \rangle_B &= \langle v_n, \lambda_n \cdot v_n \rangle_B \\ &= \lambda_n \cdot \langle v_n, v_n \rangle_B \\ &= \lambda_n \end{aligned}$$

Arnoldi Method

Approach (first eigenvector):

$$r^k = L^k\left(r^0\right) = \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n}\right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n}\right)^k + v_n\right)$$

$$v_n = \lim_{k\to\infty} \frac{r^k}{\|r^k\|_B}$$

The efficiency of convergence is determined by the separation of the largest (in absolute value) eigenvector from the next largest eigenvector:

$$\lambda_{n-1}/\lambda_n$$

Arnoldi Method

Approach (first eigenvector):

$$r^k = L^k(r^0) = \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n} \right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n} \right)^k + v_n \right)$$

$$v_n = \lim_{k \to \infty} \frac{r^k}{\|r^k\|_B}$$

If $\lambda_{n-1} = \lambda_n$, the limit:

$$\lim_{k \to \infty} \frac{r^k}{\|r^k\|_B}$$

will be some linear combination of v_{n-1} and v_n.

Since the eigenvalues are the same, this will also be an eigenvector with eigenvalue λ_n.

Arnoldi Method

Approach (first eigenvector):

$$r^k = L^k\left(r^0\right) = \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n}\right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n}\right)^k + v_n\right)$$

$$v_n = \lim_{k\to\infty} \frac{r^k}{\|r^k\|_B}$$

If $\lambda_{n-1} = -\lambda_n$, the limit:

$$\lim_{k\to\infty} \frac{r^k}{\|r^k\|_B}$$

will be some linear combination of v_{n-1} and v_n.

Since the eigenvalues are not the same, this will **not** be an eigenvector.

Arnoldi Method

Approach (first eigenvector):

$$r^k = L^k(r^0) = \alpha_n \cdot \lambda_n^k \cdot \left(\frac{\alpha_1}{\alpha_n} \cdot \left(\frac{\lambda_1}{\lambda_n} \right)^k \cdot v_1 + \cdots + \frac{\alpha_{n-1}}{\alpha_n} \cdot \left(\frac{\lambda_{n-1}}{\lambda_n} \right)^k + v_n \right)$$

$$v_n = \lim_{k \to \infty} \frac{r^k}{\|r^k\|_B}$$

If $\lambda_{n-1} =$

In our applications we will only consider self-adjoint operators $L \in \mathrm{End}(V)$ with non-negative eigenvalues.

$\Updownarrow$

The bilinear form $B \circ L \in \mathrm{Hom}(V, V^*)$ is symmetric positive semi-definite.

will be some lin

Since the eigenvalues are not the same, this will **not** be an eigenvector.

Arnoldi Method

Approach (next eigenvector):

Given eigenvectors $\{v_{m+1}, \cdots, v_n\}$ with largest (absolute) eigenvalues, we obtain the next eigenvector using a method akin to Gram-Schmidt.

Starting with a random vector r^0, we get:

$$r^0 = \alpha_1 \cdot v_1 + \cdots + \alpha_n \cdot v_n$$

Subtract off the components of $\{v_{m+1}, \cdots, v_n\}$:

$$\begin{aligned}\hat{r}^0 &= r^0 - \alpha_{m+1} \cdot v_{m+1} - \cdots - \alpha_n \cdot v_n \\ &= \alpha_1 \cdot v_1 + \cdots + \alpha_m \cdot v_m\end{aligned}$$

Since the basis is orthonormal:

$$\hat{r}^0 = r^0 - \langle r^0, v_{m+1} \rangle \cdot v_{m+1} - \cdots - \langle r^0, v_n \rangle \cdot v_n$$

Arnoldi Method

Approach (next eigenvector):

Given eigenvectors $\{v_{m+1}, \cdots, v_n\}$ with largest (absolute) eigenvalues, we obtain the next eigenvector using a method akin to Gram-Schmidt.

Starting with a random vector r^0, we get:

$$\hat{r}^0 = \alpha_1 \cdot v_1 + \cdots + \alpha_m \cdot v_m$$

Applying the k-th power of L gives:

$$\hat{r}^k = \alpha_1 \cdot \lambda_1^k + \cdots + \alpha_m \cdot \lambda_m^k \cdot v_m$$

$$= \alpha_m \cdot \lambda_m^k \cdot \left(\frac{\alpha_1}{\alpha_m} \cdot \left(\frac{\lambda_1}{\lambda_m} \right)^k \cdot v_1 + \cdots + \frac{\alpha_{m-1}}{\alpha_m} \cdot \left(\frac{\lambda_{m-1}}{\lambda_m} \right)^k + v_m \right)$$

Normalizing gives v_m (as $k \to \infty$).

Arnoldi Method

Approach (next eigenvector):

Given eigenvectors $\{v_{m+1}, \cdots, v_n\}$ with largest (absolute) eigenvalues, we obtain the next eigenvector using a method akin to Gram-Schmidt.

The implementation does **not** require an explicit expression of a matrix, just a method for evaluating the output of the linear operator on a vector.

Applyi

For numerical stability, the subtracting off of components should be done every time the operator L is applied.

In practice, this is done more efficiently using *Krylov subspaces*.

$$= \left(\alpha_m \left(\lambda_m \right) \right) \left(\frac{\lambda_{m-1}}{\lambda_m} \right)^k + v_m \Bigg)$$

Normalizing gives v_m (as $k \to \infty$).

Practical Considerations

We are often working with a symmetric positive semi-definite bilinear form, not a self-adjoint operator (e.g. the stiffness operator $S: V \to V^*$).

We are often interested in the smallest eigenvectors, not the largest.

Generalized Eigenproblem

In practice:

Given a symmetric $A \in \mathrm{Hom}(V, V^*)$, we would like to solve for a **generalized** eigen-basis $\{v_1, \ldots, v_n\}$ with **generalized** eigenvalues $\{\lambda_1, \ldots, \lambda_n\}$ such that the basis is orthonormal:

$$A(v_i) = \lambda_i \cdot B(v_i)$$

We turn this into a standard eigenvalue problem by constructing the self-adjoint operator that is the composition with the inverse of the inner-product:

$$B^{-1} \circ A \in \mathrm{End}(V)$$

Applying the Arnoldi method requires (repeatedly) solving the linear system:

$$B\left(r^{k+1}\right) = A\left(r^{k}\right).$$

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

Recall:

Assuming A is invertible, the eigenvectors of A^{-1} are the eigenvectors of A, with reciprocal eigenvalues.

$\Rightarrow$ The eigenvector of A^{-1} with largest eigenvector is the eigenvector of A with smallest eigenvalue.

$\Rightarrow$ Run the Arnoldi method using A^{-1}.

Like the generalized eigenproblem, this requires solving a linear system.

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

Recall:

Assuming A is invertible, the eigenvectors of A^{-1} are the eigenvectors of A, with recipro

> The smallest eigenvalues of A will be the reciprocals of the largest eigenvalues of A^{-1}.

$\Rightarrow$ The eigenvector of A^{-1} with largest eigenvector is the eigenvector of A with smallest eigenvalue.

$\Rightarrow$ Run the Arnoldi method using A^{-1}.

Like the generalized eigenproblem, this requires solving a linear system.

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

If A is **not** invertible, we can consider the shifted operator $A + \varepsilon \cdot \mathrm{Id}_V$.

Since A has non-negative eigenvalues, $A + \varepsilon \cdot \mathrm{Id}_V$ will have strictly positive eigenvalues and will be invertible.

$\Rightarrow$ We obtain the eigenvector of A with smallest eigenvalue by running the Arnoldi method using $\hat{A} = (A + \varepsilon \cdot \mathrm{Id}_V)^{-1}$.

Shift and Invert

In practice:

We obtain the eigenvector of A with smallest eigenvalue by running the Arnoldi method using $\hat{A} = (A + \varepsilon \cdot \mathrm{Id}_V)^{-1}$.

Note:

Assuming the first two eigenvalues of A are $\lambda_1 = 0$ and $\lambda_2 \neq 0$, the last two eigenvalues of $\hat{A}$ will be $\hat{\lambda}_n = 1/\varepsilon$ and $\hat{\lambda}_{n-1} = 1/(\lambda_2 + \varepsilon)$.

Taking the ratio gives:

$$\frac{\hat{\lambda}_n}{\hat{\lambda}_{n-1}} = \frac{\lambda_2 + \varepsilon}{\varepsilon} = 1 + \frac{\lambda_2}{\varepsilon}$$

$\Rightarrow$ For efficiency, we want ε to e as small as numerically possible.

$\hat{A} = (A + \varepsilon \cdot \mathrm{Id}_V)^{-1}$

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

The eigenvalue of the operator $\hat{A}$ associated to eigenvector v can be obtained by taking the inner-product:

$$\begin{aligned}\hat{\lambda} &= \left\langle v, \hat{A}(v) \right\rangle_B \\ &= \left\langle v, (A + \varepsilon \cdot \mathrm{Id}_V)^{-1}(v) \right\rangle_B \\ &= \left\langle v, \tfrac{1}{\lambda+\varepsilon} \cdot v \right\rangle_B \\ &= \tfrac{1}{\lambda+\varepsilon} \cdot \langle v, v \rangle_B \\ &= \tfrac{1}{\lambda+\varepsilon}\end{aligned}$$

$\hat{A} = (A + \varepsilon \cdot \mathrm{Id}_V)^{-1}$

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

The eigenvalue of the operator$\hat{A}$ associated to eigenvector v can be obtained by taking the inner-product:

$$\hat{\lambda} = \frac{1}{\lambda + \varepsilon}$$

$\Rightarrow$ We obtain the smallest eigenvalues of A by taking the reciprocal of the largest eigenvalues of $\hat{A}$ and shifting back:

$$\lambda = \frac{1}{\hat{\lambda}} - \varepsilon$$

Shift and Invert

In practice:

Given a self-adjoint $A \in \mathrm{End}(V)$ with non-negative eigenvalues, we would like to solve for the eigenvector of A with the **smallest** eigenvalue.

The eigenvalue of the operator $\hat{A}$ associated to eigenvector v can be obtained by taking the inner-product:

$$\lambda = \langle v, \hat{A}(v) \rangle_B$$

Or, we could just take the inner-product, with the evaluation of A:

$$\lambda = \langle v, A(v) \rangle_B$$

$\Rightarrow$ We obtain the smallest eigenvalues of A by taking the reciprocal of the largest eigenvalues of $\hat{A}$ and shifting back:

$$\lambda = \frac{1}{\hat{\lambda}} - \varepsilon$$

Shift and Invert

Note:

If the linear operator is represented as a symmetric positive semi-definite operator $A \in \mathrm{Hom}(V, V^*)$ we solve the shifted generalized eigenproblem:

$$\begin{aligned}\lambda \cdot v &= (B^{-1} \circ A + \varepsilon \cdot \mathrm{Id}_V)^{-1}(v) \\ &= \left(B^{-1} \circ (A + \varepsilon \cdot B)\right)^{-1}(v) \\ &= \left((A + \varepsilon \cdot B)^{-1} \circ B\right)(v)\end{aligned}$$

⇒ For the generalized eigenproblem, the Arnoldi method:

- Multiplies by B, and
- Solves the shifted system $(A + \varepsilon \cdot B)$.*

*Shifting by the identity would not make sense in any case, since $\mathrm{Id}_V \in \mathrm{End}(V)$ while $A \in \mathrm{Hom}(V, V^*)$.
In contrast, the inner-product is in the correct space, $B \in \mathrm{Hom}(V, V^*)$.