Chapter 6

Structure of Operators on Inner Product Spaces

6.1 Notes

10/11:

- Spectral decomposition of self-adjoint linear maps.
  - Can we write a map in term of the eigenvalues only?
  - Let $A : X \to X$ be linear and self-adjoint. Where $\dim X < \infty$.
  - Let A have eigenvalues $\lambda_1, \ldots, \lambda_n$ and eigenvectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$. The there is an orthonormal basis of X consisting of eigenvectors of A. An operator is self-adjoint if $A = A^*$.
  - If A is self-adjoint, then A can be written as diagonal with the eigenvalues on the diagonal with respect to some orthonormal basis of eigenvectors.
  - Let $\mathbb{F} = \mathbb{C}$.
- If there exists an orthonormal basis $u_1, \ldots, u_n$ of X such that A is triangular, then $A = UTU^*$ where U is unitary and T is upper triangular.
  - Proved with induction on $\dim X$.
  - $\dim X = 1$ is clear.
  - Assume for $\dim X = n - 1$, WTS for $\dim X = n$.
  - The subspace has a basis $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$ such that A has a diagonal form.
  - Let $u \in X$ be linearly independent of $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$.
  - Let λ be the remaining eigenvalue and u the corresponding eigenvector. Let $E = \operatorname{span}(u)$. Then make the matrix λ in the upper left corner, and block diagonal with "A_{n-1}" in the bottom right corner, zeroes everywhere else.
- **Self-adjoint** (matrix A): A linear map $A : X \to X$ where $\dim X < \infty$ such that $A = A^*$.
  - Similarly, $(Ax, y) = (x, Ay)$.
  - A self-adjoint implies all eigenvalues are real, eigenvectors corresponding to different eigenvalues are orthogonal.
    - Soug proves this.
- **Strictly positive** (operator A): A self-adjoint operator $A : X \to X$ such that $(Ax, x) > 0$ for all $x \neq 0$. *Also known as* **positive definite**.
  - Implies that all eigenvalues are strictly positive.

- **Nonnegative** (operator A): A self-adjoint operator $A : X \to X$ such that $(Ax, x) \geq 0$ for all $x \neq 0$. *Also known as* **definite**.
  - All eigenvalues are nonnegative.
- Suppose $A \geq 0$ is self-adjoint. Then there exists a unique self-adjoint $B \geq 0$ such that $B^2 = A$.
  - A self-adjoint is diagonal (wrt. some basis).
  - A positive means that all eigenvalues (diagonal entries) are positive.
  - Thus, take
$$B = \begin{pmatrix} \sqrt{\lambda_1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sqrt{\lambda_n} \end{pmatrix}$$
  - Suppose $B^2 = A$, $C^2 = A$. Then we have an orthonormal basis corresponding to B and an orthonormal basis corresponding to C. It follows that $B^2 = C^2 = A$. Write B^2x and C^2x in terms of their bases; will necessitate that the bases are the same.

10/13:

- If we get yes/no questions, we don't have to justify.
- Cauchy-Schwarz inequality:
$$|(\mathbf{x}, \mathbf{y})| \leq \|\mathbf{x}\|\|\mathbf{y}\|$$
  - Real spaces, V vs. $(\cdot, \cdot)$ inner product.
  - Proof:
$$\begin{aligned} 0 &\leq \|\mathbf{x} + t\mathbf{y}\|^2 \\ &= t^2\|\mathbf{y}^2\| + 2t(\mathbf{x}, \mathbf{y}) + \|\mathbf{x}\|^2 \end{aligned}$$
    Thus, the discriminant must be less than zero (because the whole polynomial is positive, so the discriminant [the opposite of the x^0 term of the factored form of the polynomial] must be less than zero so the polynomial doesn't get dragged down to negative values):
$$(\mathbf{x}, \mathbf{y})^2 - \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \leq 0$$
    Taking square roots of both sides proves the desired inequality.
- Recall that if $A^* = A$, then all eigenvalues are real and all eigenvectors of distinct eigenvalues are orthogonal to each other.
- **Normal** (matrix): A matrix N such that $N^*N = NN^*$.
  - Examples: Diagonal, self-adjoint, and unitary operators are all normal.
- Any normal operator in a complex vector spae has an orthonormal set of eigenvectors, e.g., $N = UDU^*$.
  - Proof: N is upper triangular wrt. some basis (because all matrices are). WTS any normal upper triangular matrix is diagonal. Done by induction on the dimension of N from $n = 2$.
  - Assume the claim for every $(n-1) \times (n-1)$ normal upper triangular matrix.
  - Let
$$N = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ 0 & & & \\ \vdots & & N_1 & \\ 0 & & & \end{pmatrix}$$
    (we know every normal matrix can be written in this upper triangular form)

  - Then just compute NN^* and N^*N. Knowing they have to be equal, we have that $a_{12} = \cdots = a_{1n} = 0$.
  - We can also prove from the above (block diagonal multiplication) that N_1 is normal. Thus, it's diagonal, too. Therefore, the whole thing is diagonal.

- N is normal if and only if $\|N\mathbf{x}\| = \|N^*\mathbf{x}\|$.
  - Proof: $(N\mathbf{x}, N\mathbf{y}) = (N^*N\mathbf{x}, \mathbf{y}) = (NN^*\mathbf{x}, \mathbf{y}) = (N^*\mathbf{x}, N^*\mathbf{y})$. This is equivalent to the desired condition.

- If A is nonnegative and $(A\mathbf{e}_k, \mathbf{e}_k) = a_{kk}$, then

$$\sum_{i,j=1}^{n} a_{ij}\mathbf{x}_i\mathbf{x}_j$$

- **Positive definite** (matrix): An $n \times n$ self-adjoint matrix such that $(A\mathbf{x}, \mathbf{x}) > 0$ for all $\mathbf{x} \in X$.

- Let $A : X \to Y$, $\dim X = \dim Y$. Then AA^* is positive semidefinite. And there exists a unique square root $R = \sqrt{A^*A}$.
  - Proof: $(A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) = \|A\mathbf{x}\|^2 \geq 0$.

- **Modulus** (of A): The matrix $|A| = \sqrt{A^*A}$.

- Check $\||A|\mathbf{x}\| = \|A\mathbf{x}\|$.

$$\||A|\mathbf{x}\|^2 = (|A|\mathbf{x}, |A|\mathbf{x}) = (|A|^*|A|\mathbf{x}, \mathbf{x}) = (A^*A\mathbf{x}, \mathbf{x}) = (A\mathbf{x}, A\mathbf{x}) = \|A\mathbf{x}\|^2$$

- Let $A : X \to X$ be a linear operator. Then $A = U|A|$ where U is unitary.

- Look at singular matrices.

10/15: - Recall that if $A : X \to Y$, we have that A^*A is semidefinite, positive, and self adjoint.
  - Thus, there exists a unique matrix $R = \sqrt{A^*A} \geq 0$, which we define to be $|A| = \sqrt{A^*A}$.

- Polar form of a matrix:

$$A = U|A|$$

  - This may not be unique!
  - Proof: Suppose $A\mathbf{x} = U(|A|\mathbf{x})$. $A\mathbf{x} \in \operatorname{range} A$, and $|A|\mathbf{x} \in \operatorname{range}(|A|)$. $\mathbf{x} \in \operatorname{range}(|A|)$ implies that there exists $\mathbf{v} \in X$ such that $x = |A|\mathbf{v}$.
  - Define $U\mathbf{x} = A\mathbf{x}$. U is a well-defined linear map.
  - $\|U_0\mathbf{x}\| = \|A\mathbf{x}\| = \||A|\mathbf{v}\| = \|\mathbf{x}\|$.
  - U is an isometry.
  - $\operatorname{range}|A| \to X$.
  - Use $\ker A = \ker|A| = (\operatorname{range} A)^\perp$ to extend U_0 to U: $U = U_0 + U_1$.

- **Singular values** (of a matrix): The eigenvalues of $|A|$.
  - So if $\lambda_1, \dots, \lambda_n$ are the eigenvalues of A^*A, the singular values of A are $\sqrt{\lambda_1}, \dots, \sqrt{\lambda_n}$.

- Let $A : X \to Y$ be a linear map.
  - Let $\sigma_1, \dots, \sigma_n$ be the signular values of A. Then $\sigma_1, \dots, \sigma_n > 0$.
  - Additionally, if $\mathbf{v}_1, \dots, \mathbf{v}_n$ is an orthonormal basis of eigenvectors of A^*A, then the list of n vectors $\mathbf{w}_1, \dots, \mathbf{w}_n$ defined by $\mathbf{w}_k = 1/\sigma_k A\mathbf{v}_k$ for each $k = 1, \dots, n$ is orthonormal.

- Proof:

$$(\mathbf{w}_k, \mathbf{w}_j) = \frac{1}{\sigma_k\sigma_k}(A\mathbf{v}_k, A\mathbf{v}_j) = \frac{1}{\sigma_k\sigma_j} = \frac{1}{\sigma_k\sigma_j}(A^*A\mathbf{v}_k, \mathbf{v}_j) = \frac{\sigma_k^2}{\sigma_k\sigma_j}(\mathbf{v}_k, \mathbf{v}_j) = 0$$

and

$$\|\mathbf{w}_k\| = \frac{1}{\sigma_k}\|A\mathbf{v}_k\| = \frac{1}{\sigma_k}\||A|\mathbf{v}_k\| = 1$$

– Schmidt decomposition of A:

$$A\mathbf{x} = \sum^r \sigma_k(\mathbf{x}, \mathbf{v}_k)\mathbf{w}_k$$

- This is because $\mathbf{x} = \sum(\mathbf{x}, \mathbf{v}_k)\mathbf{v}_k$, so by the above,

$$A\mathbf{x} = \sum^n(\mathbf{x}, \mathbf{v}_k)A\mathbf{v}_k = \sum^r \sigma_k(\mathbf{x}, \mathbf{v}_k)\mathbf{w}_k$$

- **Operator norm**: $\|A\| = \max\{\|A\mathbf{x}\| : \|\mathbf{x}\| \leq 1\}$.
- Properties of the operator norm:
  - $\|A\mathbf{x}\| \leq \|A\|\|\mathbf{x}\|$.
  - $\|\alpha A\| = |\alpha|\|A\|$.
  - $\|A + B\| \leq \|A\| + \|B\|$.
  - $\|A\| \geq 0$.
  - $\|A\| = 0$ iff $A = 0$.
- **Frobenius norm**: The norm $\|A\|_2^2 = \operatorname{tr}(A^*A)$.
- The operator norm is always less than or equal to the Frobenius norm.
- If $A : \mathbb{F}^n \to \mathbb{F}^n$, then $A = W\Sigma V^*$ where σ is a diagonal matrix of nonzero singular values.
- The operator norm of A is the largest of the singular values.
- An orthogonal matrix can be decomposed to a block-diagonal matrix of rotations.

10/18:
- Soug tests what he teaches and doesn't give super tricky questions.
- Structure of orthogonal matrices.
- **Orthogonal** (matrix): A unitary matrix U with all elements real and $|\det U| = 1$.
- Theorem: Let U be an orthogonal operator on $\mathbb{R}^n$ such that $\det U = 1$. Then there exists an orthonormal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that with respect to this basis,

$$U = \begin{pmatrix} R_{\phi_1} & & & \mathbf{0} \\ & \ddots & & \\ & & \mathbb{R}_{\phi_k} & \\ \mathbf{0} & & & I_{n-2k} \end{pmatrix}$$

  where each R_{ϕ_i} is a 2×2 rotation matrix.
  - If you are in $\mathbb{R}^7$ for example, you would be able to express U as a composition of at most 3 rotation maps and the identity map.
  - Each rotation map acts on two orthonormal vectors.
  - Proof: $P(\lambda)$ is the n-degree characteristic polynomial $\det(U - \lambda I) = 0$. The eigenvalues are the roots of it.

- $p(\lambda) = 0$ if and only if $p(\bar{\lambda}) = 0$.
  - $\lambda \in \mathbb{C}$ is an eigenvalue with eigenvector $\mathbf{u} \neq 0$ iff $U\mathbf{u} = \lambda\mathbf{u}$ and $U\bar{\mathbf{u}} = \bar{\lambda}\bar{\mathbf{u}}$.
- Recall that U unitary implies $|\lambda| = 1$.
  - Proof[1]: $\|U\mathbf{x}\| = \|\mathbf{x}\|$ and $U\mathbf{x} = \lambda\mathbf{x}$. Thus,

$$\|U\mathbf{x}\| = \|\lambda\mathbf{x}\| = |\lambda|\|\mathbf{x}\| = \|\mathbf{x}\|$$

    and since $\mathbf{x} \neq 0$, we can divide by $\|\mathbf{x}\|$, so $|\lambda| = 1$.
- Let $\mathbf{u} = \operatorname{Re}\mathbf{u} + \operatorname{Im}\mathbf{u}$.
- It follows that we may define

$$\mathbf{x} = \operatorname{Re}\mathbf{u} = \frac{\mathbf{u} + \bar{\mathbf{u}}}{2} \qquad\qquad \mathbf{y} = \operatorname{Im}\mathbf{u} = \frac{\mathbf{u} - \bar{\mathbf{u}}}{2}$$

- Thus, $\mathbf{u} = \mathbf{x} + i\mathbf{y}$ and $\bar{\mathbf{u}} = \mathbf{x} - i\mathbf{y}$.
- Since $U\mathbf{x} = \frac{U\mathbf{u}+U\bar{\mathbf{u}}}{2} = \frac{\lambda\mathbf{u}+\bar{\lambda}\bar{\mathbf{u}}}{2}$, $U\mathbf{y} = \operatorname{Im}(\lambda\mathbf{u}) = \operatorname{Re}(\lambda\mathbf{u})$.
- Since $|\lambda| = 1$, $\lambda = \mathrm{e}^{i\alpha}$ and $\bar{\lambda} = \mathrm{e}^{-i\alpha}$.
- It follows that $U\mathbf{x} = (\cos\alpha)\mathbf{x} - (\sin\alpha)\mathbf{y}$ and $U\mathbf{y} = (\cos\alpha)\mathbf{y}) + (\sin\alpha)\mathbf{x}$.
- Thus, since $U\mathbf{x} = \operatorname{Re}\lambda\mathbf{u}$, we have that

$$\begin{aligned}\lambda\mathbf{u} &= (\cos\alpha + i\sin\alpha)(\mathbf{x} + i\mathbf{y}) \\ &= (\cos\alpha)\mathbf{x} - (\sin\alpha)\mathbf{y} + i[(\cos\alpha)\mathbf{y} + (\sin\alpha)\mathbf{x}]\end{aligned}$$

- If E_λ is a 2 dimensional space spanned by $\mathbf{x}$ and $\mathbf{y}$ and invariant by U. Thus, any block of the desired matrix leaves its desired sub-block invariant.
- We also know that the eigenvectors of a unitary matrix corresponding to different eigenvalues are orthogonal.
- Thus, $\|\mathbf{x}\| = \|\mathbf{y}\| = \sqrt{2}/2\|\mathbf{u}\|$, $\mathbf{x} \perp \mathbf{y}$.
- Let $\mathbf{x}, \mathbf{y}$ complete the theorem to form a basis of $\mathbb{R}^n$.
- It will follow that

$$U = \begin{pmatrix} R_\alpha & \mathbf{0} \\ \mathbf{0} & U_1 \end{pmatrix}$$

  where U_1 is orthogonal, and we may repeat the process.

6.2 Chapter 6: Structure of Operators on Inner Product Spaces

From Treil (2017).

10/24:

- Theorem 6.1.1: Let $A : X \to X$ be an operator acting in a complex inner product space. Then there exists an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ of X such that the matrix of A in this basis is upper triangular. In other words, any $n \times n$ matrix A can be represented as $A = UTU^*$, where U is unitary and T is upper-triangular.

- Theorem 6.1.2: Let $A : X \to X$ be an operator acting on a real inner product space. Suppose that all eigenvalues of A are real. Then there exists an orthonormal basis $\mathbf{u}_1, \dots, \mathbf{u}_n$ in X such that the matrix of A in this basis is upper triangular. In other words, any real $n \times n$ matrix A with all real eigenvalues can be represented as $T = UTU^* = UTU^T$, where U is orthogonal and T is a real upper-triangular matrix.

[1] This would be a good exam question.

- Theorem 6.2.1: Let $A = A^*$ be a self-adjoint operator in an inner product space X (the space can be complex or real). Then all eigenvalues of A are real and there exists an orthonormal basis of eigenvectors of A in X.

  Equivalently (see Theorem 6.2.2), A can be represented as $A = UDU^*$ where U is a unitary matrix and D is a diagonal matrix with real entries. Moreover, if A is real, U can be chosen to be real, i.e., orthogonal.

- Proposition 6.2.3: Let $A = A^*$ be a self-adjoint operator and let $\lambda, \mathbf{u}, \mu, \mathbf{v}$ be such that $A\mathbf{u} = \lambda\mathbf{u}$ and $A\mathbf{v} = \mu\mathbf{v}$. Then if $\lambda \neq \mu$, $\mathbf{u} \perp \mathbf{v}$.

- Since complex multiplication is commutative,

  $$D^*D = DD^*$$

  for every diagonal matrix D.

  - It follows that $A^*A = AA^*$ if the matrix of A in some orthonormal basis is diagonal.

- Theorem 6.2.4: Any normal operator N in a complex vector space has an orthonormal basis of eigenvectors.

  Equivalently, any matrix N satisfying $N^*N = NN^*$ can be represented as $N = UDU^*$ where U is unitary and D is diagonal.

- Proposition 6.2.5: An operator $N : X \to X$ is normal iff

  $$\|N\mathbf{x}\| = \|N^*\mathbf{x}\|$$

  for all $\mathbf{x} \in X$.

- **Hermitian square** (of A): The matrix A^*A.

- **Modulus** (of A): The unique positive semidefinite square root $\sqrt{A^*A}$.

- Proposition 6.3.3: For a linear operator $A : X \to Y$,

  $$\||A|\mathbf{x}\| = \|A\mathbf{x}\|$$

- Corollary 6.3.4: $\ker A = \ker |A|$.

- Theorem 6.3.5: Let $A : X \to X$ be an operator (square matrix). Then A can be represented as

  $$A = U|A|$$

  where U is a unitary operator.

- **Singular value** (of A): An eigenvalue of $|A|$.

  - A positive square root of an operator of A^*A.

- Proposition 6.3.6: Let $\sigma_1, \dots, \sigma_n$ be the singular values of A, ordered such that $\sigma_1, \dots, \sigma_r$ are the nonzero singular values, and let $\mathbf{v}_1, \dots, \mathbf{v}_n$ be an orthonormal basis of eigenvectors of A^*A. Then the system

  $$\mathbf{w}_k = \frac{1}{\sigma_k} A\mathbf{v}_k$$

  for $k = 1, \dots, r$ is orthonormal.

- **Schmidt decomposition** (of A): The decompositions

$$A = \sum_{k=1}^{r} \sigma_k \mathbf{w}_k \mathbf{v}_k^*$$

and

$$A\mathbf{x} = \sum_{k=1}^{r} \sigma_k (\mathbf{x}, \mathbf{v}_k) \mathbf{w}_k$$

  - Note that these can be verified by plugging $\mathbf{x} = \mathbf{v}_j$ for each $j = 1, \dots, n$ into the latter equation.

10/25: • Lemma 6.3.7: A can be represented as the Schmidt decomposition

$$A = \sum_{k=1}^{r} \sigma_k \mathbf{w}_k \mathbf{v}_k^*$$

where $\sigma_k > 0$ for any orthonormal systems $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$.

- Corollary 6.3.8: Let $A = \sum_{k=1}^{r} \sigma_k \mathbf{w}_k \mathbf{v}_k^*$ be a Schmidt decomposition of A. Then

$$A^* = \sum_{k=1}^{r} \sigma_k \mathbf{v}_k \mathbf{w}_k^*$$

is a Schmidt decomposition of A^*.

- **Reduced singular value decomposition** (of A): The decomposition

$$A = \tilde{W} \tilde{\Sigma} \tilde{V}^*$$

where $A : \mathbb{F}^n \to \mathbb{F}^m$ has the Schmidt decomposition $A = \sum_{k=1}^{r} \sigma_k \mathbf{w}_k \mathbf{v}_k^*$, $\tilde{\Sigma} = \operatorname{diag}\{\sigma_1, \dots, \sigma_r\}$, and $\tilde{V}, \tilde{W}$ are matrices with columns $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$, respectively. *Also known as* **compact singular value decomposition**.

  - Note that $\tilde{V}$ is an $n \times r$ matrix, $\tilde{\Sigma}$ is an $r \times r$ matrix, and $\tilde{W}$ is an $m \times r$ matrix.
  - Since $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$ are orthonormal, $\tilde{V}, \tilde{W}$ are isometries.

- Note that $r = \operatorname{rank} A$ (see Problem 6.3.1).
  - It follows that if A is invertible, then $m = n = r$, so $\tilde{V}, \tilde{W}$ are unitary and $\tilde{\Sigma}$ is an invertible diagonal matrix.

- However, A need not be invertible for us to get a representation similar to $A = \tilde{W} \tilde{\Sigma} \tilde{V}^*$.
  - Complete $\mathbf{v}_1, \dots, \mathbf{v}_r$ and $\mathbf{w}_1, \dots, \mathbf{w}_r$ to bases of $\mathbb{F}^n$ and $\mathbb{F}^m$, respectively.
  - Then we get the following.

- **Singular value decomposition** (of A): The decomposition

$$A = W \Sigma V^*$$

where $V \in M_{n \times n}^{\mathbb{F}}$ and $W \in M_{m \times m}^{\mathbb{F}}$ are unitary matrices with columns $\mathbf{v}_1, \dots, \mathbf{v}_n$ and $\mathbf{w}_1, \dots, \mathbf{w}_n$, respectively, and $\Sigma \in M_{m \times n}^{\mathbb{R}^+}$ is a "diagonal" matrix such that

$$\Sigma_{j,k} = \begin{cases} \sigma_k & j = k \leq r \\ 0 & \text{otherwise} \end{cases}$$

- Notice that if $A = W\Sigma V^*$, then
$$A^*A = (W\Sigma V^*)^*(W\Sigma V^*) = V\Sigma^*W^*W\Sigma V^* = V\Sigma^2 V^*$$
proving that the singular values of A, squared, are the eigenvalues of A^*A.
- If A is invertible, the reduced SVD is the matrix form of the Schmidt decomposition is the SVD.
- If $A = W\Sigma V^*$ is $n \times n$, then
$$A = (\underbrace{WV^*}_{U})(\underbrace{V\Sigma V^*}_{|A|})$$
is a polar decomposition of A.
- Consider the unit ball $B = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| \leq 1\}$.
  - We want to describe $A(B)$, i.e., the image of the unit ball under A.
  - Let $\mathbf{x} = (x_1, \dots, x_n)^T$ and let $\mathbf{y} = (y_1, \dots, y_n)^T$. If $A = \operatorname{diag}\{\sigma_1, \dots, \sigma_n\}$, we have $\mathbf{y} \in A(B)$ iff $\mathbf{y} = A\mathbf{x}$ where $\mathbf{x} \in B$ iff
$$\sum_{k=1}^{n} \frac{y_k^2}{\sigma_k^2} = \sum_{k=1}^{n} x_k^2 = \|\mathbf{x}\|^2 \leq 1$$
  - Thus, $A(B)$ is an ellipsoid with half-axes $\sigma_1, \dots, \sigma_n$.
  - In the more general case, if $A = W\Sigma V^*$, then since V^* is unitary, $V^*(B) = B$. $\Sigma V^*(B) = \Sigma(B)$ is thus by the above an ellipsoid in range Σ with half-axes $\sigma_1, \dots, \sigma_r$. Thus, since isometries don't change geometry, $W(\Sigma(B))$ is also an ellipsoid with the same half-axes, but in range A.
- Conclusion: The image $A(B)$ of the closed unit ball B is an ellipsoid in range A with half-axes $\sigma_1, \dots, \sigma_r$, where r is the number of nonzero singular values, i.e., the rank of A.
- Finding the maximum of $\|A\mathbf{x}\|$ for $\mathbf{x} \in B$.
  - For a diagonal matrix Σ with nonnegative entries, the maximum is clearly the maximal diagonal entry: In this case if s_1 is the maximal diagonal entry, then since
$$\Sigma\mathbf{x} = \sum_{k=1}^{r} s_k x_k \mathbf{e}_k$$
we have that
$$\|A\mathbf{x}\|^2 = \sum_{k=1}^{r} s_k^2 |x_k|^2 \leq s_1^2 \sum_{k=1}^{r} |x_k|^2 = s_1^2 \cdot \|\mathbf{x}\|^2$$
  - We get the following by a similar logic to before.
- Conclusion: The maximum of $\|A\mathbf{x}\|$ on the unit ball B is the maximal singular value of A.
- **Operator norm** (of A): The following quantity. *Denoted by* $\|\boldsymbol{A}\|$. *Given by*
$$\|A\| = \max\{\|A\mathbf{x}\| : \mathbf{x} \in X, \|\mathbf{x}\| \leq 1\}$$
  - $\|A\|$ clearly satisfies the four properties of a norm.
  - Additionally,
$$\|A\mathbf{x}\| \leq \|A\| \cdot \|\mathbf{x}\|$$
  - Alternate definition: The operator norm $\|A\|$ is the smallest number $C \geq 0$ such that $\|A\mathbf{x}\| \leq C\|\mathbf{x}\|$.

- **Frobenius norm**: The following norm. *Also known as* **Hilbert-Schmidt norm**. *Denoted by* $\|\boldsymbol{A}\|_{\mathbf{2}}$. *Given by*
$$\|A\|_2^2 = \operatorname{tr}(A^*A)$$
  - If we let $s_1, \dots, s_n$ be the singular values of A and let s_1 be the largest value, then we have
$$\|A\|^2 = s_1^2 \leq \sum_{k=1}^{n} s_k^2 = \operatorname{tr}(A^*A) = \|A\|_2^2$$
- Conclusion: The operator norm of a matrix cannot be more than its Frobenius norm.
- Suppose we want to solve $A\mathbf{x} = \mathbf{b}$ where A is invertible, but there is some (experimental) error $\Delta\mathbf{b}$ in $\mathbf{b}$. Then we are really solving for an approximate solution $\mathbf{x} + \Delta\mathbf{x}$ to the equation
$$A(\mathbf{x} + \Delta\mathbf{x}) = \mathbf{b} + \Delta\mathbf{b}$$
  - It follows since A is invertible that $\mathbf{x} = A^{-1}\mathbf{b}$ and $\Delta\mathbf{x} = A^{-1}\Delta\mathbf{b}$.
  - To estimate the relative error $\|\Delta\mathbf{x}\|/\|\mathbf{x}\|$ in the solution in comparison with the relative error $\|\Delta\mathbf{b}\|/\|\mathbf{b}\|$ in the data, use
$$\frac{\|\Delta\mathbf{x}\|}{\|\mathbf{x}\|} = \frac{\left\|A^{-1}\Delta\mathbf{b}\right\|}{\|\mathbf{b}\|}\frac{\|A\mathbf{x}\|}{\|\mathbf{x}\|} \leq \frac{\left\|A^{-1}\right\| \cdot \|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}\frac{\|A\| \cdot \|\mathbf{x}\|}{\|\mathbf{x}\|} = \left\|A^{-1}\right\| \cdot \|A\| \cdot \frac{\|\Delta\mathbf{b}\|}{\|\mathbf{b}\|}$$
- **Condition number** (of A): The following quantity. *Given by*
$$\|A\| \cdot \left\|A^{-1}\right\|$$
  - If s_1 is the largest singular value of A and s_n is the smallest, then
$$\|A\| \cdot \left\|A^{-1}\right\| = s_1 \cdot \frac{1}{s_n} = \frac{s_1}{s_n}$$
- **Well-conditioned** (matrix): A matrix the condition number of which is not "too big."
- **Ill-conditioned** (matrix): A matrix that is not well-conditioned.
- Theorem 6.5.1: Let U be an orthogonal operator on $\mathbb{R}^n$ and let $\det U = 1$. Then there exists an orthonormal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that the matrix of U in this basis has the block diagonal form
$$\begin{pmatrix} R_{\varphi_1} & & & 0 \\ & \ddots & & \\ & & R_{\varphi_k} & \\ 0 & & & I_{n-2k} \end{pmatrix}$$
where each R_{φ_j} is a two-dimensional rotation
$$R_{\varphi_j} = \begin{pmatrix} \cos\varphi_j & -\sin\varphi_j \\ \sin\varphi_j & \cos\varphi_j \end{pmatrix}$$
and I_{n-2k} represents the $(n-2k) \times (n-2k)$ identity matrix.
  - Alternate interpretation: Any rotation in $\mathbb{R}^n$ can be represented as a composition of at most $n/2$ commuting planar rotations.

- Theorem 6.5.2: Let U be an orthogonal operator on $\mathbb{R}^n$ and let $\det U = -1$. Then there exists an orthonormal basis $\mathbf{v}_1, \dots, \mathbf{v}_n$ such that the matrix of U in this basis has block diagonal form
$$\begin{pmatrix} R_{\varphi_1} & & & & 0 \\ & \ddots & & & \\ & & R_{\varphi_k} & & \\ & & & I_r & \\ 0 & & & & -1 \end{pmatrix}$$
where $r = n - 2k - 1$ and each R_{φ_j} is a two-dimensional rotation
$$R_{\varphi_j} = \begin{pmatrix} \cos\varphi_j & -\sin\varphi_j \\ \sin\varphi_j & \cos\varphi_j \end{pmatrix}$$
- Corollary: An orthogonal 2×2 matrix U with determinant -1 is always a reflection.
- Theorem 6.5.3: Any rotation U (i.e., any orthogonal transformation U with $\det U = 1$) can be represented as a product of at most $n(n-1)/2$ elementary rotations.
- Consider the following orthonormal bases of $\mathbb{R}^2$.

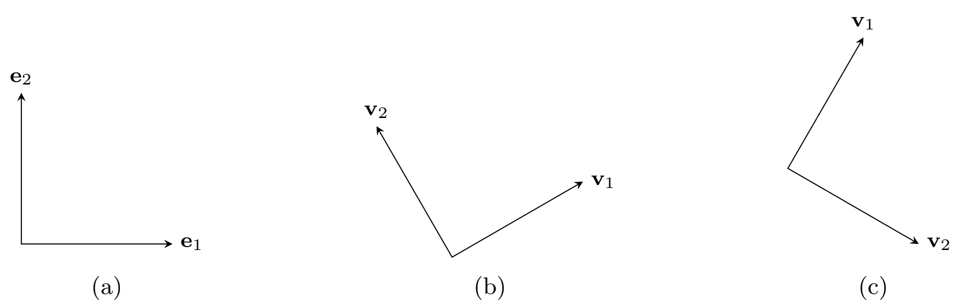

Figure 6.1: Orientation in $\mathbb{R}^2$.

  - Notice that a rotation will get you from the standard basis (a) to basis (b), but not from the standard basis (a) to basis (c).
  - This is the motivation for defining orientation.
  - More formally, we know that there is a unique linear transformation U such that $U\mathbf{e}_k = \mathbf{v}_k$ for each $k = 1, 2$. In particular, the matrix of U with respect to the standard basis is orthogonal with columns $\mathbf{v}_1, \mathbf{v}_2$.
  - By Theorems 6.5.1 and 6.5.2, if $\det U = 1$, then U is a rotation, and if $\det U = -1$, then U is not a rotation.
- **Similarly oriented** (bases $\mathcal{A}, \mathcal{B}$): Two bases $\mathcal{A}, \mathcal{B}$ of a real vector space such that the change of coordinates matrix $[I]_{\mathcal{BA}}$ has a positive determinant.
- **Differently oriented** (bases $\mathcal{A}, \mathcal{B}$): Two bases $\mathcal{A}, \mathcal{B}$ of a real vector space that are not similarly oriented (i.e., $[I]_{\mathcal{BA}}$ has a negative determinant).
- We usually let the standard basis of $\mathbb{R}^n$ have a **positive orientation**.
  - In an abstract vector space, we need only fix a basis and declare its orientation to be positive.
- **Continuously transformable** (bases $\mathcal{A}, \mathcal{B}$): Two bases $\mathcal{A}, \mathcal{B}$ such that $\mathcal{A} = \{\mathbf{a}_1, \dots, \mathbf{a}_n\}$ can be continuously transformed to a basis $\mathcal{B} = \{\mathbf{b}_1, \dots, \mathbf{b}_n\}$. In particular, there exists a **continuous family of bases** $\mathcal{V}(t) = \{\mathbf{v}_1(t), \dots, \mathbf{v}_n(t)\}$, $t \in [a, b]$, such that
$$\mathbf{v}_k(a) = \mathbf{a}_k \qquad\qquad \mathbf{v}_k(b) = \mathbf{b}_k$$
for each $k = 1, \dots, n$.

- **Continuous family of bases**: A family of bases $\mathcal{V}(t) = \{\mathbf{v}_1(t), \ldots, \mathbf{v}_n(t)\}$, $t \in [a, b]$, such that the vector-functions $\mathbf{v}_k(t)$ are continuous (their coordinates in some bases are continuous functions) and the system $\mathbf{v}_1(t), \ldots, \mathbf{v}_n(t)$ is a basis for all $t \in [a, b]$.
- Theorem 6.6.1: Two bases $\mathcal{A} = \{\mathbf{a}_1, \ldots, \mathbf{a}_n\}$ and $\mathcal{B} = \{\mathbf{b}_1, \ldots, \mathbf{b}_n\}$ have the same orientation if and only if one of the bases can be continuously transformed to the other.