Condensed Linear Algebra II notes*

Salman A. Faris

10 March 2019

1 Basic Linear Algebra recap

Let $T : V \to W$ be linear.

Definition 1.1. (Kernel). The *kernel* of T (also known as *null space* of T) is the set of all elements $\mathbf{v} \in V$ such that $T\mathbf{v} = \mathbf{0}$.

$$\mathrm{Ker}(T) = \{\mathbf{v} \in V \mid T\mathbf{v} = \mathbf{0}\} \tag{1.1}$$

Layman 1.1. The kernel is everything that is mapped to zero by the transformation/function.

Definition 1.2. (Image). The *image* of T (also known as *range* of T) is the set of all values that T can take as its argument varies over V, i.e.

$$\mathrm{Im}(T) = \{T\mathbf{v} \in W \mid \mathbf{v} \in V\} \tag{1.2}$$

$\boxed{\text{One can see that if } \mathbf{v} \in \mathrm{Im}(f)\text{, then } \mathbf{v} = f(\textit{junk}) \text{ for some } \textit{junk} \text{ in } V.}$

Layman 1.2. The image is everything that comes out of the function.

*based on notes written by Dr. Thomas Bothner and Dr. Mahir Hadzic in module 4CCM114A of King's College London, Spring 2019 and notes by Dr. Kwok-Wing Tsoi in module 5CCM222A.

2 Diagonalization Theorem

2.1 Characteristic Polynomial

Theorem 2.1. Let $A \in \mathbb{C}^{n\times n}$ and let $\chi(\lambda) = \det(A - \lambda I)$ be its characteristic polynomial. Then $\chi(\lambda)$ is a polynomial of degree n. Moreover, $\chi(\lambda)$ has the form:

$$\chi(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}\operatorname{trace}(A)\lambda^{n-1} + \ldots + \det(A) \tag{2.1}$$

Recipe 2.1. (χ_2). The characteristic polynomial of a 2×2 matrix. When $n = 2$:

$$\chi_2(\lambda) = \lambda^2 - \operatorname{trace}(A)\lambda + \det(A) \tag{2.2}$$

Recipe 2.2. (χ_3). The characteristic polynomial of a 3×3 matrix. When $n = 3$:

$$\chi_3(\lambda) = -\lambda^3 + \operatorname{trace}(A)\lambda^2 - (A_{11} + A_{22} + A_{33})\lambda + \det(A) \tag{2.3}$$

2.2 Eigenvectors and Eigenvalues

There are a lot of recipes to find eigenvectors from corresponding eigenvalues.

Let A be an $n \times n$ matrix and that $\lambda \in \sigma(A)$ is an eigenvalue of A and that its corresponding eigenvector is denoted $\mathbf{v}$. Then, $A\mathbf{v} = \mu\mathbf{v} \implies (A - \mu I)\mathbf{v} = 0$

Recipe 2.3. (Reduced-echelon e-vector). One can find the eigenvector by the following process:

1. Substitute λ into $A - \mu I$ and compute this subtraction.
2. Reduce $A - \lambda I$ into its row-reduced echelon form.
3. Compute $(A - \lambda I)\mathbf{v} = 0$ and solve for components of $\mathbf{v}$.

Trick 2.1. (Hadzic). In solving imaginary eigenvalues, in particular, when $\lambda = \pm i$ in Recipe 2.3, one can eliminate it by noticing that:

$$\boxed{i^2 - 1 = (i + 1)(i - 1) = -2} \tag{2.4}$$

Recipe 2.4. (Brute force). One can find the eigenvector by the following process:

1. Substitute λ into $A\mathbf{v} = \mu\mathbf{v}$.
2. Solve for components of $\mathbf{v}$.

Despite this algorithm being shorter, its actual computation takes more time compared to the reduced-echelon method.

There's a handy method to find eigenvectors relating to 2×2 matrices.

Recipe 2.5. (Watts). Let A be a 2×2 matrix, $\lambda \in \sigma(A)$ and $\mathbf{v}$ the corresponding eigenvector. Then the recipe is as follows:

1. Let $\mathbf{v} = (1, a)$ for some $a \in \mathbb{F}$.
2. Solve $A\mathbf{v} = \lambda\mathbf{v}$ with the eigenvector $(1, a)$, in particular, solve for a.

We can generalise this method for any 2×2 matrix.

Theorem 2.2. The minimal polynomial of a matrix A divides any polynomial $f(x)$ satisfied by A. In other words:

$$\boxed{m_A(x) \mid f(x) \iff f(A) = 0} \tag{2.5}$$

Proof. The " $\implies$ " direction is trivial. Suppose $m_A(x) \mid f(x)$, then $f(x) = m_A(x)q(x)$ and it immediately follows that $f(A) = 0$ by the definition of minimal polynomial.

Now we proof in this " $\impliedby$ " direction. Suppose $f(A) = 0$. Let $f(x) = m_A(x)q(x) + r(x)$ where $m_A(x)$ is the minimal polynomial of A and by the Division Algorithm Lemma, $\deg(r(x)) < \deg(m(x))$ or $r(x) = 0$. Then let $x = A$, and we have $f(A) = m_A(A)q(A) + r(A)$. But by our assumption and the Cayley-Hamilton theorem on the minimal polynomial, we have that:

$$0 = 0 + r(A) \implies r(A) = 0$$

By the minimality of the degree of $m(x)$, the only possibility is that $r(x) = 0$. So we are able to rewrite $f(x) = m_A(x)q(x)$, and thus $m_A(x) \mid f(x)$ as required. $\square$

3 Inner Products Spaces

3.1 Classical Gram-Schmidt process (CGS)

Definition 3.1. We define the **projection operator** to be:

$$\mathrm{proj}_{\mathbf{u}}(\mathbf{v}) = \frac{\langle \mathbf{v}, \mathbf{u} \rangle}{\langle \mathbf{u}, \mathbf{u} \rangle}\mathbf{u}$$

where $\langle \mathbf{u}, \mathbf{v} \rangle$ is the *inner product* of the vectors $\mathbf{u}$ and $\mathbf{v}$, i.e. $\langle \mathbf{u}, \mathbf{v} \rangle = \mathbf{u}^*\mathbf{v}$ for vectors on some general vector space $\mathbb{F}^n$. This operator projects the vector $\mathbf{v}$ orthogonally onto the line spanned by the vector $\mathbf{u}$. If $\mathbf{u} = \mathbf{0}$, we define $\mathrm{proj}_0(\mathbf{v}) := 0$, i.e. the projection map proj_0 is the zero map, sending every vector to the zero vector.

Theorem 3.1. (Gram-Schmidt Orthogonalization & Orthonormalization).
The CGS is a process **to orthonormalise a set of vectors in an inner product space**. It takes a finite, linearly independent set $S = \{\mathbf{u}_1, \ldots, \mathbf{u}_k\}$ for $k \leq n$ and generates an orthogonal set $S' = \{\mathbf{o}_1, \ldots, \mathbf{o}_n\}$ that spans the same k-dimensional subspace of $\mathbb{F}^n$ as S. The Gram-Schmidt process then commence as follows:

$$\begin{aligned}
&\mathbf{o}_1 = \mathbf{u}_1, && \mathbf{e}_1 = \frac{\mathbf{o}_1}{||\mathbf{o}_1||} \\
&\mathbf{o}_2 = \mathbf{u}_2 - \mathrm{proj}_{\mathbf{o}_1}(\mathbf{u}_2), && \mathbf{e}_2 = \frac{\mathbf{o}_2}{||\mathbf{o}_2||} \\
&\mathbf{o}_3 = \mathbf{u}_3 - \mathrm{proj}_{\mathbf{o}_1}(\mathbf{u}_3) - \mathrm{proj}_{\mathbf{o}_2}(\mathbf{u}_3), && \mathbf{e}_3 = \frac{\mathbf{o}_3}{||\mathbf{o}_3||} \\
&\mathbf{o}_4 = \mathbf{u}_4 - \mathrm{proj}_{\mathbf{o}_1}(\mathbf{u}_4) - \mathrm{proj}_{\mathbf{o}_2}(\mathbf{u}_4) - \mathrm{proj}_{\mathbf{o}_3}(\mathbf{u}_4), && \mathbf{e}_4 = \frac{\mathbf{o}_4}{||\mathbf{o}_4||} \\
&\quad \vdots && \quad \vdots \\
&\mathbf{o}_k = \mathbf{u}_k - \sum_{j=1}^{k-1} \mathrm{proj}_{\mathbf{o}_j}(\mathbf{u}_k), && \mathbf{e}_k = \frac{\mathbf{o}_k}{||\mathbf{o}_k||}
\end{aligned}$$

Remark 3.1. (Important remark on the CGS). The Gram-Schmidt process consists of repeated applications of so called **orthogonal projections**, a procedure which arises naturally in the solution of the following approximation problem:

Suppose W is a subspace of an inner product space V and $\beta \in V$ be arbitrary. The problem is to find a vector $\alpha \in W$ such that $||\beta - \alpha||$ is a small as possible.

Definition 3.2. Let W be a subspace of an inner product space V and $\beta \in V$. A **best approximation** to β by vectors in W is a vector $\alpha \in W$ such that:

$$\boxed{||\beta - \alpha|| \leq ||\beta - \gamma|| \quad \forall \gamma \in W} \tag{3.1}$$

Theorem 3.2. (Best approximation). Let W be a subspace of an inner product space V and $\beta \in V$.

- The vector $\alpha \in W$ is a best approximation to β by vectors in W **if and only if** $\beta - \alpha$ is orthogonal to every vector in W.
- If a best approximation to β by vectors in W exists, then it is **unique**.
- If W is finite-dimensional with ONB $\{\alpha_1, \dots, \alpha_n\}$, then,

$$\alpha = \sum_{k=1}^{n} \frac{\langle \beta, \alpha_k \rangle}{||\alpha_k||^2} \alpha_k$$

  is the *unique* best approximation to β by vectors in W.

Definition 3.3. Let V be an inner product space and S any set of vectors in V. The **orthogonal complement** of S is the set $S^\perp$ of all vectors in V which are orthogonal to every vector in S.

Note that $V^\perp = \{0\}$ and $\{0\}^\perp = V$; also $S^\perp$ is always a subspace, since it contains 0 and whenever $\alpha, \beta \in S^\perp, c \in F$, then $\forall \gamma \in S$,

$$\langle c\alpha + \beta, \gamma \rangle = c\langle \alpha, \gamma \rangle + \langle \beta, \gamma \rangle = 0$$

so, $c\alpha + \beta \in S^\perp$. The above also let us rephrase Theorem 3.2: α is the only vector in W such that $\beta - \alpha \in W^\perp$

Definition 3.4. (Orthogonal projection of β on W).
Whenever the vector α in Theorem 3.2 exists, we call it the **orthogonal projection of β on W**.

Definition 3.5. (Orthogonal projection of V on W).
If all vectors in V have an orthogonal projection on W, the mapping that assigns to each vector in V its orthogonal projection on W is called the **orthogonal projection of V on W**. By Theorem 3.2, **the orthogonal projection of an inner product space on a finite-dimensional subspace always exists.**

Corollary 3.1. (Orthogonal projection of V on $W^\perp$).
Let V be an inner product space, W a finite-dimension subspace, and E the orthogonal projection of V on W. Then the mapping,

$$\beta \longmapsto \beta - E\beta$$

is the orthogonal projection of V on $W^\perp$.

Theorem 3.3. (Decomposition theorem - super simplified).
Let $E : V \to V$ be a linear map. If E is *idempotent*, then:

$$\boxed{V = \operatorname{Im}(E) \oplus \operatorname{Ker}(E)} \tag{3.2}$$

Definition 3.6. Let $W_1, W_2 \subset V$ be subspaces of vector space V. We say that V is a **direct sum** of W_1 and W_2 or

$$V = W_1 \oplus W_2 \tag{3.3}$$

if the following two conditions hold:

1. $W_1 + W_2 := \{\alpha + \beta : \alpha \in W_1, \beta \in W_2\} = V$
   i.e. *The span of W_1 and W_2 is all of V.*

2. $W_1 \cap W_2 = \{\mathbf{0}\}$
   i.e. *W_1 and W_2 are independent.*

Theorem 3.4. (Decomposition theorem - complete definition).
Let W be a finite-dimensional subspace of an inner product space V and E be the orthogonal projection of V on W. Then E is an **idempotent** linear transformation of V onto W (i.e. $E^2 = E$), $W^\perp$ is the kernel of E, and

$$\boxed{V = W \oplus W^\perp} \tag{3.4}$$

Corollary 3.2. Under the conditions of Theorem 3.4, $I - E$ is the orthogonal projection of V on $W^\perp$. It is an idempotent linear transformation of V onto $W^\perp$ with kernel W.

Corollary 3.3. (Bessel's inequality)
Let $\{\alpha_1, \alpha_2, \ldots, \alpha_n\}$ be an orthogonal set of non-zero vectors in an inner product space V. For any $\beta \in V$, we have:

$$\boxed{\sum_{k=1}^{n} \frac{|\langle \beta, \alpha_k \rangle|^2}{||\alpha_k||^2} \leq ||\beta||^2}$$

and equality holds if and only if

$$\beta = \sum_{k=1}^{n} \frac{\langle \beta, \alpha_k \rangle}{||\alpha_k||^2} \alpha_k$$

3.2 Linear Functionals and Adjoints

Let V be any inner product space and $\beta \in V$ some fixed vector. Clearly:

$$f_\beta(\alpha) := \langle \alpha, \beta \rangle, \quad \alpha \in V$$

is a linear functional on V. Turns out, if V is finite-dimensional, every linear functional on V arises in this way from some β.

Theorem 3.5. (Riesz representation).
Let V be a finite-dimensional inner product space, and f a linear functional on V. Then, there exists a unique vector $\beta \in V$ such that $f(\alpha) = \langle \alpha, \beta \rangle$ for all $\alpha \in V$. In two lines, we can rephrase this as the following:

If V f-d-IPS and f linear functional on V,

$$\exists \text{ unique } \beta \in V, \text{ s.t. } f(\alpha) = \langle \alpha, \beta \rangle \quad \forall\, \alpha \in V. \tag{3.5}$$

Theorem 3.6. (Existence of adjoint). For any linear operator T on a finite-dimensional inner product space V, there exists a unique linear operator T^* on V s.t.

$$\boxed{\langle T\alpha, \beta \rangle = \langle \alpha, T^*\beta \rangle \quad \forall \alpha, \beta \in V} \tag{3.6}$$

Theorem 3.7. Let V f-d-IPS with ONB $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$. Let T be a linear operator on V and A be the matrix of T in the basis $\mathcal{B}$. Then:

$$\boxed{A_{kj} = \langle T\alpha_j, \alpha_k \rangle} \tag{3.7}$$

Corollary 3.4. Let V be a f-d-IPS and T a linear operator on V. In any ONB for V, the matrix of T^* is the conjugate transpose of the matrix of T.

Definition 3.7. Let T be a linear operator on an IPS V. We say that T has an adjoint on V if there exists a linear operator T^* on V s.t.

$$\langle T\alpha, \beta \rangle = \langle \alpha, T^*\beta \rangle \quad \forall \alpha, \beta \in V$$

In light of Theorem 3.6 (Adjoint existence), we know then that:

$$\boxed{\text{Every linear operator on a f-d-IPS } V \text{ has an adjoint on } V.}$$

This is not always true if $\dim(V) = \infty$. But it any case, there is at most one such operator T^* and where it exists, we call it the adjoint of T.

Theorem 3.8. Let V be a f-d-IPS. If T and U are linear operators on V and $c \in F$:

1. $(T + U)^* = T^* + U^*$

2. $(cT)^* = \overline{c}\, T^*$
3. $(TU)^* = U^*T^*$
4. $(T^*)^* = T$

Definition 3.8. Let V be an IPS. A linear operator T s.t. $T = T^*$ is called self-adjoint or Hermitian.

If $\mathcal{B}$ is an ONB for V, then

$$[T^*]_{\mathcal{B}} = [T]_{\mathcal{B}}^*$$

so T is self-adjoint *if and only if* its matrix in every ONB is a Hermitian matrix.

Theorem 3.9. Let V be a f-d-IPS and T linear operator on V. Then:

$$\boxed{\mathrm{Im}(T^*) = \mathrm{Ker}(T)^{\perp}}$$

Theorem 3.10. Let V be a f-d-IPS and T linear operator on V. Then:

$$\boxed{T \text{ invertible } \implies T^* \text{ invertible } \textbf{and } (T^*)^{-1} = (T^{-1})^*}$$

Theorem 3.11. Let V be a f-d-IPS and E *idempotent* linear operator on V i.e. $E^2 = E$. Then:

$$\boxed{E \text{ self-adjoint } \iff E \text{ normal}}$$

3.3 Unitary Operators

Definition 3.9. Let V and W be IPS over the same field, and T a linear transformation from V into W. We say T preserves inner products if:

$$\boxed{\langle T\alpha, T\beta\rangle_W = \langle \alpha, \beta\rangle_V \ , \quad \forall \alpha, \beta \in V} \tag{3.8}$$

An isomorphism of V onto W is a vector space isomorphism $T : V \to W$ which also preserves inner products.

If T preserves inner products, then:

$$\boxed{||T\alpha|| = ||\alpha|| \implies T \text{ necessarily non-singular} \implies T \text{ invertible}} \tag{3.9}$$

Thus, an isomorphism from V onto W is simply *a linear transformation which preserves inner products.*

Definition 3.10. If an isomorphism of V onto W exists, we will simply say that V and W are isomorphic.

Theorem 3.12. (TFAE Unitary Operators). Let V and W be f-d-IPS over the same field with $\dim V = \dim W$ (same dimension). If $T : V \to W$ linear, then the following are equivalent:

(i) T preserves inner products

(ii) T is an (IPS) isomorphism

(iii) T carries every ONB for V onto an ONB for W.

(iv) T carries some ONB for V onto an ONB for W.

Corollary 3.5. Let V and W be f-d-IPS over the same field. Then:

$$\boxed{V \text{ and } W \text{ isomorphic} \iff \dim V = \dim W}$$

Theorem 3.13. Let V and W be IPS over the same field and $T : V \to W$ linear. Then:

$$\boxed{T \text{ preserves inner products} \iff ||T\alpha|| = ||\alpha||,\ \forall \alpha \in V}$$

Definition 3.11. A unitary operator on an IPS is an isomorphism of the space onto itself.

Important things from this definition! Note that the set of all unitary operators on an IPS is a *group over composition.* Moreover, Theorem 3.12 tells us the following:

$$\boxed{U \text{ unitary} \iff \langle U\alpha, U\beta\rangle = \langle \alpha, \beta\rangle,\ \forall \alpha, \beta \in V}$$

We can see this by this process

$$\begin{aligned} U \text{ unitary} &\Leftrightarrow U \text{ isomoprhism} \\ &\Leftrightarrow U \text{ preserves inner products} \\ &\Leftrightarrow \text{for some ONB } \{\alpha_1, \dots, \alpha_n\}, \{U\alpha_1, \dots, U\alpha_n\} \text{ is an ONB.} \\ &\Leftrightarrow \langle U\alpha, U\beta\rangle = \langle \alpha, \beta\rangle,\ \forall \alpha, \beta \in V \end{aligned}$$

Theorem 3.14. Let U be a linear operator on an IPS V. Then:

$$\boxed{U \text{ unitary} \iff \text{the adjoint } U^* \text{ of } U \text{ exists } \textbf{and } UU^* = U^*U = I}$$

Definition 3.12. (Unitary matrix).

$$\boxed{\text{A matrix } A \in \mathbb{C}^{n\times n} \text{ is called unitary if } A^*A = I.}$$

Theorem 3.15. Let V be a f-d-IPS and U be a linear operator on V. Then:

$$\boxed{U \text{ unitary} \iff \text{the matrix of } U \text{ in some (every) ONB is a unitary matrix.}}$$

Remark 3.2. For $A \in \mathbb{C}^{n\times n}$,

$$\boxed{A \text{ unitary } \textit{simply means } (A^*A)_{jk} = \delta_{jk}}$$

or equivalently:

$$\boxed{\text{A unitary} \implies \sum_{r=1}^{n} \overline{A}_{rj} A_{rk} = \delta_{jk}}$$

In other words, the columns of A forms an orthonormal set of column vectors in $\mathbb{C}^n$ with respect to the standard inner product.

Theorem 3.16. (TFAE Unitary Matrices). Let $U \in \mathbb{C}^{n\times n}$. Then, the following are equivalent:

1. U is unitary i.e. $U^*U = I$,
2. U is invertible and satisfies $U^* = U^{-1}$,
3. The rows of U forms an ONB for $\mathbb{C}^n$,
4. The columns of U forms an ONB for $\mathbb{C}^n$.
5. $||Uv|| = ||v||$ for all $v \in \mathbb{C}^n$
6. $\langle Uv, Uw\rangle = \langle v, w\rangle$ for all $v, w \in \mathbb{C}^n$

Remark 3.3. Let U and U' be unitary matrices. Then:

$$\overline{U}, U^T, \overline{U}^*, UU' \text{ all are unitary.}$$

Definition 3.13. (Orthogonal matrix).

$$\boxed{\text{A real or complex } n \times n \text{ matrix } A \text{ is called orthogonal if } A^T A = I}$$

Remark 3.4. A real orthogonal matrix is unitary

Remark 3.5. A unitary matrix is orthogonal *if and only if* its entries are real.

Definition 3.14. Let $A, B \in \mathbb{C}^{n\times n}$. Then,

$$\boxed{B \text{ is unitarily equivalent to } A \text{ if } \exists P \in \mathbb{C}^{n\times n} \text{ unitary s.t. } B = P^{-1}AP}$$

and

$$\boxed{B \text{ is orthogonally equivalent to } A \text{ if } \exists P \in \mathbb{R}^{n\times n} \text{ orthogonal s.t. } B = P^{-1}AP}$$

3.4 Normal operators

Motivation: If $T : V \to V$ and V an IPS, under which conditions does V have an ONB consisiting of *eigenvectors* for T?

Definition 3.15. Let V be an IPS and $T : V \to V$ linear. Then:

T normal if it commutes with its adjoint i.e. $TT^* = T^*T$

Notice here we assume implicitly that T^* exists!

Remark 3.6. Here are some interesting remarks:

1. Any self-adjoint operator is normal
2. Any unitary operator is normal
3. Any scalar multiple of a normal operator is normal

Danger 3.1. *Sums and products of normal operators* are in general not normal.

Theorem 3.17. Let V be an IPS and T a self-adjoint linear operator on V. Then:

Each eigenvalue of T is **real** (every eigenvalue of matrix of T is real), and **eigenvectors** of T associated with **distinct** eigenvalues are **orthogonal**.

But note that the Theorem above says nothing about the **existence** of eigenvalues or eigenvectors!

Theorem 3.18. On a f-d-IPS V of positive dimension i.e. $\dim V > 0$,

Every self-adjoint operator has a (non-zero) eigenvector.

Remark 3.7. Note the following remarks:

1. Existence of $X \neq 0$ s.t. $AX = cX$ does not rely at all on the fact that A is Hermitian. This is *only* used in the case of a real IPS, then it tells us that all eigenvalues of A are real **and** all entries of X are real as well.
2. The **characteristic polynomial of a Hermitian matrix has real coefficients**, in spite of the fact that the matrix may not have real entries.
3. Finite dimensional vector spaces are **really needed**.

Theorem 3.19. (Invariant). Let V be a f-d-IPS and $T : V \to V$ linear. Suppose $W \subseteq V$ a subspace of V which is invariant under T (i.e. $T\alpha \in W, \forall \alpha \in W$). Then $W^\perp$ is invariant under T^*.

Theorem 3.20. (Spectral Theorem). Let V be a f-d-IPS, and $T : V \to V$ self-adjoint. Then:

$$\boxed{\text{There is an ONB for } V\text{, each vector of which is an eigenvector for } T.}$$

Corollary 3.6. Let $A \in \mathbb{C}^{n\times n}$ be Hermitian (self-adjoint). Then, $\exists$ unitary $P \in \mathbb{C}^{n\times n}$ s.t. $P^{-1}AP$ diagonal (A unitarily equivalent to a diagonal matrix).

Corollary 3.7. If $A \in \mathbb{R}^{n\times n}$ is symmetric, $\exists$ real orthogonal $P \in \mathbb{R}^{n\times n}$ s.t. $P^{-1}AP$ is diagonal.

Remark 3.8. If V is a f-d-**Real**-IPS and $T : V \to V$ linear, then V has an ONB of eigenvectors for T *if and only if* T is self-adjoint. *There is no such equivalence for complex symmetric matrices!*

Theorem 3.21. Let V be an IPS and $T : V \to V$ normal. Suppose $\alpha \in V$. Then:

$$\boxed{\alpha \text{ is an eigenvector for } T \text{ with eigenvalue } c \text{ if and only if } \alpha \text{ is an eigenvector for } T^* \text{ with eigenvalue } \bar{c}.}$$

Definition 3.16. (Normal matrix).

$$\boxed{\text{A matrix } A \in \mathbb{C}^{n\times n} \text{ is normal if } AA^* = A^*A}$$

Theorem 3.22. Let $A \in \mathbb{C}^{n\times n}$ be a normal matrix. Then:

$$\boxed{||Av|| = ||A^*v|| \text{ for all } v \in \mathbb{C}^n}$$

Proof.

$$||Av||^2 = \langle Av, Av\rangle = \langle v, A^*Av\rangle = \langle v, AA^*v\rangle = \langle A^*v, A^*v\rangle = ||A^*v||^2$$

$\square$

Theorem 3.23. If A is normal, then A and A^* have the same eigenvectors.

Theorem 3.24. Let V be a f-d-IPS, $T : V \to V$ linear and $\mathcal{B}$ an ONB for V. Suppose that $A_T = [T]_{\mathcal{B}}$ upper triangular. Then:

$$\boxed{T \text{ normal} \iff A_T \text{ diagonal}}$$

Theorem 3.25. (Schur triangulization). Let V be a f-d-**Complex**-IPS and $T : V \to V$ linear. Then:

$$\boxed{\text{There is an ONB for } V \text{ in which the matrix of } T \text{ is upper-triangular.}}$$

Corollary 3.8. For every $A \in \mathbb{C}^{n\times n}$:

$$\boxed{\exists \text{ unitary matrix } U \text{ s.t. } U^{-1}AU \text{ upper-triangular.}}$$

3.5 Spectral Theorem (Main theorem)

Theorem 3.26. (Spectral theorem). Let V be a f-d-**Complex**-IPS and that $T : V \to V$ **normal**. Then:

$$V \text{ has an ONB consisting of eigenvectors for } T.$$

Definition 3.17. $A \in \mathbb{C}^{n\times n}$ is *unitarily diagonalizable* if $\exists\, U$ unitary such that U^*AU diagonal.

Theorem 3.27. ($\mathbb{C}$-Analogue of the Spectral Theorem). Any normal matrix is diagonalizable.

$$A \in \mathbb{C}^{n\times n} \textit{ unitarily diagonalizable} \iff A \text{ normal}$$

Definition 3.18. $A \in \mathbb{R}^{n\times n}$ is *orthogonally diagonalizable* if $\exists\, P$ orthogonal such that $P^T AP$ diagonal.

Theorem 3.28. ($\mathbb{R}$-Analogue of the Spectral Theorem). Any symmetric matrix is diagonalizable.

$$A \in \mathbb{R}^{n\times n} \textit{ orthogonally diagonalizable} \iff A \text{ symmetric}$$

Corollary 3.9. $A \in \mathbb{C}^{n\times n}$ is self-adjoint $\iff$ A unitarily diagonalizable with only **real** eigenvalues.

Layman 3.1. A "layman" interpretation [**mathumich**] of the Spectral Theorem is that suppose A is normal/symmetric, then we can write A as $P^{-1}DP$ with P consisting of orthonormal vectors as its columns.

4 Operators on Inner Product Spaces

4.1 Forms on Inner Product Spaces

Motivation: If $T : V \to V$ linear on a f-d-IPS V, the function $f : V \times V \to F$ defined by:

$$f(\alpha, \beta) := \langle T\alpha, \beta \rangle$$

may be regarded as kind of a substitute for T. This is due to *many questions concerning T are equivalent to questions concerning f*. In particular, f determines T. With an ONB $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ we find at once that for $A = [T]_{\mathcal{B}}$,

$$A_{jk} = f(\alpha_k, \alpha_j)$$

Definition 4.1. (Sesqui-linear form). A sesqui-linear form on a real or complex vector space V is a function f on $V \times V$ with values in the scalar field s.t.

1. $f(c\alpha + \beta, \gamma) = cf(\alpha, \gamma) + f(\beta, \gamma)$
2. $f(\alpha, c\beta + \gamma) = \bar{c}f(\alpha, \beta) + f(\alpha, \gamma)$

for all $\alpha, \beta, \gamma \in V$ and $c \in F$.

Remark 4.1. (Bilinear form). If $F = \mathbb{R}$, we call f in Definition 4.1 a bilinear form.

Remark 4.2. If f and g are forms and $c \in F$, then $cf + g$ is also a form! This tells us that the set of all forms on V is a subspace of the vector space of all scalar-valued functions on $V \times V$.

Theorem 4.1. (Representation). Let V be a f-d-IPS and f a form on V. Then:

$$\boxed{\exists \text{ unique linear operator } T : V \to V \text{ s.t. } f(\alpha, \beta) = \langle T\alpha, \beta \rangle \ \forall \alpha, \beta \in V}$$

Moreover, the map $f \mapsto T$ is an isomorphism of the space of forms onto $\mathcal{L}(V, V)$.

Corollary 4.1. The equation:

$$\langle f, g \rangle := \operatorname{trace}(T_f T_g^*) \tag{4.1}$$

defines an inner product on the space of forms with the property that:

$$\langle f, g \rangle = \sum_{j,k=1}^{n} f(\alpha_k, \alpha_j)\overline{g(\alpha_k, \alpha_j)} \tag{4.2}$$

for every ONB $\{\alpha_1, \ldots, \alpha_n\}$ of V.

Definition 4.2. (Matrix of forms in ordered basis). If f is a form on $V \times V$ and $\mathcal{B} = \{\alpha_1, \ldots, \alpha_n\}$ an arbitrary basis of V, the matrix $A \in F^{n\times n}$ with entries

$$A_{jk} = f(\alpha_k, \alpha_j)$$

is called the matrix of f in the ordered basis $\mathcal{B}$.

Remark 4.3. If $\mathcal{B}$ is an ONB, then the above A is also the matrix of T_f, but in general this is not the case!

Theorem 4.2. (Triangulization). Let f be a form on a f-d-**Complex**-IPS V and A_f matrix of f. Then:

$$\boxed{\exists \text{ an ONB for } V \text{ in which } A_f \text{ is upper-triangular}}$$

Definition 4.3. (Hermitian forms). A form f on a real or complex vector space V is:

$$\boxed{\text{Hermitian if } f(\alpha,\beta) = \overline{f(\beta,\alpha)},\ \forall \alpha,\beta \in V}$$

Remark 4.4. If $T : V \to V$ linear on a f-d-IPS V and f is the (standard representation) form $f(\alpha, \beta) = \langle T\alpha, \beta\rangle$, then observe that:

$$\overline{f(\beta,\alpha)} = \langle \alpha, T\beta\rangle = \langle T^*\alpha, \beta\rangle$$

so $\boxed{f \text{ is Hermitian } \textit{if and only if } T \text{ self adjoint}}$. Also, when f is Hermitian, then $f(\alpha, \alpha) \in \mathbb{R}$,$\forall \alpha \in V$ and on *complex spaces this property characterizes Hermitian forms*. We will see this in the next theorem.

Theorem 4.3. (Hermitian characterization of forms). Let V be a **COMPLEX** vector space and f a form on V s.t. $f(\alpha, \alpha) \in \mathbb{R}$,$\forall \alpha \in V$. Then f is Hermitian.

Corollary 4.2. Let $T : V \to V$ on a f-d-**Complex**-IPS V. Then:

$$\boxed{T \text{ self-adjoint } \iff \langle T\alpha, \alpha\rangle \in \mathbb{R} \text{ ,} \forall \alpha \in V}$$

Theorem 4.4. (Principal axis theorem). For every Hermitian form f on a f-d-IPS V:

$$\boxed{\exists \text{ an ONB of } V \text{ in which } f \text{ is represented by a diagonal matrix with } \textbf{real entries}.}$$

Corollary 4.3. Under the assumptions of the last theorem:

$$f\left(\sum_{j=1}^{n} x_j\alpha_j, \sum_{k=1}^{n} y_k\alpha_k\right) = \sum_{j=1}^{n} c_j x_j \overline{y_j}$$

4.2 Positive Forms

Definition 4.4. A form f on a real or complex vector space V is:

1. Non-negative if f Hermitian and $f(\alpha, \alpha) \geq 0$ for every $\alpha \in V$.
2. Positive if f Hermitian and $f(\alpha, \alpha) > 0$ for all $\alpha \neq 0$.

Remark 4.5. We have the following remarks:

1. A positive form is simply an inner product.
2. A non-negative form satisfies all the properties of an inner product **except that some non-zero vectors may be *orthogonal* to themselves.**

Theorem 4.5. Let $F = \mathbb{R}$ or $F = \mathbb{C}$ and $A \in F^{n\times n}$. Define the function g to be $\boxed{g(X, Y) := Y^*AX}$. Then:

$$\boxed{g \text{ positive form on } F^{n\times 1} \iff \exists \text{ invertible } P \in F^{n\times n} \text{ s.t. } A = P^*P}$$

Think $A = P^*P$ like the square root of a matrix.

In practice, it is hard to verify that a given $A \in F^{n\times n}$ satisfies the criteria for positivity which we have given this far. *Still*, the last Theorem tells us that if g is positive, then $\det A > 0$ because:

$$\det A = \det(P^*P) = \det(P^*)\det(P) = \overline{\det(P)}\det(P) = |\det P|^2 > 0$$

However, the fact that $\det A > 0$ is by **no means sufficient to guarantee that g is positive!** We need to look at more determinants.

Definition 4.5. (Principal minors). Let $A = F^{n\times n}$. The principal minors of A are the scalars $\Delta_k(A)$ defined by:

$$\boxed{\Delta_k(A) := \det\begin{pmatrix} A_{11} & \dots & A_{1k} \\ \vdots & \ddots & \vdots \\ A_{k1} & \dots & A_{kk} \end{pmatrix}, \quad k = 1, 2, \dots, n} \tag{4.3}$$

Lemma 4.1. (TFAE Principal Minors). Let $A \in F^{n\times n}$ be an invertible matrix. The following are equivalent:

1. $\exists$ upper-trianguler matrix P with $P_{kk} = 1 (1 \leq k \leq n)$ s.t. the matrix $B = AP$ is lower-triangular.
2. The principal minors of A are all different from zero i.e.

$$\Delta_k(A) \neq 0, \ \forall k = 1, \dots, n$$

Theorem 4.6. Let f be a form on a f-d vector space V and let A be the matrix of f in a basis $\mathcal{B}$. Then:

$$\boxed{f \text{ positive form} \iff A = A^* \text{ and } \Delta_k(A) > 0.}$$

Remark 4.6. We proceed with asking the following question: *What is it that characterizes the matrices which represent positive forms?*

1. (Complex Story). If f is a form on a **complex vector space** and A its matrix in some basis, then:

   $$\boxed{f \text{ positive} \iff A = A^* \textbf{ and } X^*AX > 0,\ \forall \textbf{ complex } X \neq 0.}$$

   It follows from Theorem 4.3 that the condition $A = A^*$ is redundant, i.e. $X^*AX > 0,\ \forall X \neq 0 \implies A = A^*$.

2. (Real Story). If f is a form on a **real vector space** and A its matrix in some basis, then:

   $$\boxed{f \text{ positive} \iff A = A^T \textbf{ and } X^TAX > 0,\ \forall \textbf{ real } X \neq 0.}$$

   But note that this condition does not imply that $A = A^T$. We need a third condition.

3. If $A = A^T$ and $X^TAX > 0$ for all **real** $X \neq 0$, then:

   $$\boxed{X^*AX > 0 \text{ for all } \textbf{complex } X \neq 0} \text{ holds as well.}$$

Definition 4.6. (Positive matrix).

$$\boxed{A \text{ positive matrix if } A \in \mathbb{C}^{n\times n} \text{ satisfies } A = A^* \text{ and } X^*AX > 0, \forall X \in \mathbb{C}^n \backslash \{0\}}$$

Let f be a form on a real or complex vector space and A_f be its matrix. From Definition 4.6, we can see that Remark 4.6 showed that in either *real* or *complex* story:

$$f \text{ is } \textbf{positive} \iff A_f \text{ in every basis is a } \textbf{positive matrix}$$

Definition 4.7. A linear operator T on a f-d-IPS V is:

1. Non-negative if $T = T^*$ **and** $\langle T\alpha, \alpha\rangle \geq 0,\ \forall \alpha \in V$.
2. Positive if $T = T^*$ **and** $\langle T\alpha, \alpha\rangle > 0,\ \forall \alpha \neq 0$.

We are now ready to summarize all our findings.

Summary 4.1. (TFAE for $A \in \mathbb{C}^{n\times n}$). If $A \in \mathbb{C}^{n\times n}$, the following are equivalent:

1. A is positive, i.e. $\sum_{j,k=1}^{n} A_{kj} x_j \overline{x_k} > 0$ whenever $x_1, \ldots, x_n \in \mathbb{C}$, not all zero.
2. $\langle X, Y\rangle := Y^* AX$ is an inner product on $\mathbb{C}^{n\times 1}$.
3. The linear operator $X \mapsto AX$ is positive, relative to the standard inner product on $\mathbb{C}^{n\times 1}$.
4. $\exists$ invertible $P \in \mathbb{C}^{n\times n}$ s.t. $A = P^* P$.
5. $A = A^*$ and the principal minors A are positive (i.e. $\Delta_k(A) > 0$).

Summary 4.2. (TFAE for $A \in \mathbb{R}^{n\times n}$). If $A \in \mathbb{R}^{n\times n}$, then the following are equivalent:

1. $A = A^T$ and $\sum_{j,k=1}^{n} A_{kj} x_j x_k > 0$ whenever $x_1, \ldots, x_n \in \mathbb{R}$ not all zero.
2. $\langle X, Y\rangle := Y^T AX$ is an inner product on $\mathbb{R}^{n\times 1}$.
3. The linear operator $X \mapsto AX$ is positive, relative to the standard inner product on $\mathbb{R}^{n\times 1}$.
4. $\exists$ invertible $P \in \mathbb{R}^{n\times n}$ s.t. $A = P^T P$

4.3 Applications to Conic Sections

A quadratic curve in $(x, y) \in \mathbb{R}^2$ is an algebraic curve of the form:

$$\boxed{ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0} \tag{4.4}$$

with coefficients $a, b, c, d, e, f \in \mathbb{R}$.

The three classical conic sections i.e. ellipses, hyperbolas and parabolas are examples of such quadratic curves **and** using the principal axis theorem, these are in fact the only *generic* quadratic curves in $(x, y) \in \mathbb{R}^2$.

First observation: Given a curve of the form

$$\boxed{ax^2 + 2bxy + cy^2 + 2dx + 2ey + f = 0}$$

we can transform it into canonical form:

$$\boxed{a\tilde{x}^2 + 2b\tilde{x}\tilde{y} + c\tilde{y}^2 = g} \tag{4.5}$$

using a simple translation $\tilde{x} = x - x_0$, $\tilde{y} = y - y_0$ provided $ac - b^2 \neq 0$ for some choice of x_0 and y_0.

Second observation: For a quadratic curve in canonical form (*note we remove the tilde because we know it can always be reduced to this canonical form*):

$$ax^2 + 2bxy + cy^2 = g \iff (x, y)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = g \tag{4.6}$$

Then:

$$\boxed{\exists \text{ orthogonal matrix } P \in \mathbb{R}^{2\times 2} \text{ s.t. } \lambda_1(x')^2 + \lambda_2(y')^2 = g}$$

where $\begin{pmatrix} x \\ y \end{pmatrix} = P\begin{pmatrix} x' \\ y' \end{pmatrix}$ and λ_1, λ_2 eigenvalues of $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

Indeed, $f(x, y) := (x, y)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$ defines a Hermitian quadratic form on $\mathbb{R}^2 \times \mathbb{R}^2$, thus the principal axis theorem **guarantees existence of an ONB** in which f is represented by a diagonal matrix. This immediately yields the above.

Remark 4.7. The two change of variables correspond geometrically to translation and rotation of coordinate matrices.

1. $(x, y) \mapsto (\tilde{x}, \tilde{y})$ corresponds to translation.
2. $(\tilde{x}, \tilde{y}) \mapsto (x', y')$ corresponds to rotation. In particular, this rotation is governed by the P matrix.