Quadratic Forms with Applications

A Thesis
Submitted to the African University of Science and Technology
Abuja-Nigeria
in partial fulfilment of the requirements for

MASTER DEGREE IN PURE AND APPLIED MATHEMATICS

By
OLULANA OLUSEUN MORENIKEJI

Supervisor:

Dr Guy Degla

African University of Science and Technology
www.aust.edu.ng
P.M.B 681, Garki, Abuja F.C.T
Nigeria.

2013 M.Sc Project

Contents

Preface

The scope of Quadratic Form Theory is historically wide although it usually appears almost as an afterthought when needed to solve a variety of problems such as the classification of Hessian matrices in finite dimensional Calculus [1], [2], [3], the finding of invariants that fully describe the equivalence class of a given form in Algebraic Geometry and Number Theory [4], the use of Rayleigh-Ristz methods for finding eigenvalues of real symmetric matrices in Linear Algebra [5], [6], the second order optimality conditions in Optimization Theory [1], [2], [3], the Sturm comparison criteria and the Sturm-Liouville Boundary Value Problems in Differential Equations [5], the kinetic energy or the Hamiltonian in Mechanics [8], etc...

In Advanced Mathematics, Quadratic Forms occupy a central place in various branches including Number Theory, Algebra, Group Theory (orthogonal groups) [7], [4], [29], Calculus of Variations and Optimal Control Theory (the second variational problem)[9], Operator Theory [8], Differential Geometry (Riemannian metrics and fundamental forms) [11], Morse Theory (Morse lemma) [12], [13], Differential Topology (intersection forms of four-manifolds), and Lie Theory (the Killing form) [14], [15].

In this dissertation, our aim is to review the Theory of Quadratic Forms on Euclidean and Hermitian spaces, to give an idea of its generalization to Hilbert spaces and to mention some common applications including the Linear Regression, the Mean Square Approximation, the Rayleigh-Ristz method and the Lax-Milgram Theorem (the bounded as well as the unbounded cases).

Introduction

A quadratic form over a field $\mathbb{K}$ ($\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) in finitely many indeterminates $x_1, \ldots, x_n$ is a homogeneous polynomial of degree 2 in $\mathbb{K}[x_1, \ldots, x_n]$, unless it is identically zero. The main property of real quadratic forms in the finite dimensional case is that every real quadratic form is orthogonally similar (i.e., can be transformed by an orthogonal change of the indeterminates considered as coordinates) to a quadratic form which is the sum of multiples of squares of the indeterminates [16].
In fact a quadratic form in a finite set of indeterminates over $\mathbb{K}$, as a homogeneous quadratic polynomial in the indeterminates with coefficients in $\mathbb{K}$, can be studied by means of matrices because any such a quadratic form Q can be expressed as $Q(X) = X^T \mathbf{A} X$, where X is a column vector with the indeterminates as elements and $\mathbf{A}$ a symmetric matrix over $\mathbb{K}$. Thus it is the quadratic form associated with the symmetric bilinear form defined from $\mathbb{K}^n \times \mathbb{K}^n$ to $\mathbb{K}$ by

$$f(X, Y) \;=\; X^T \mathbf{A} Y\,; \qquad X,\, Y \in \mathbb{K}^n\,,$$

and this gives rise to a duality.

A change of coordinates for a quadratic form Q over $\mathbb{K}$ can be made by the relation/equation $X = PY$ where the column vector X contains the original indeterminates and P is a nonsingular matrix over $\mathbb{K}$. Under the transformation $X = PY$, the matrix $\mathbf{A}$ is transformed into $P^T \mathbf{A} P = P^{-1} \mathbf{A} P$ and the quadratic form becomes $q(Y) := Q(PY) = Y^T (P^T \mathbf{A} P) Y$. Knowing from basic Linear Algebra that if the symmetric matrix $\mathbf{A}$ is real, then $\mathbb{R}^n$ admits an orthonormal basis of eigenvectors $(e_1, \ldots, e_n)$ of $\mathbf{A}$ with respective real eigenvalues $(\lambda_1, \ldots, \lambda_n)$, and taking P as the matrix of whose respective column vectors are the eigenvectors, we have that

$$q(Y) \;:=\; Q(PY) \;=\; \sum_{k=1}^{n} \lambda_k y_k^2 \qquad \text{for every column} - \text{vector} \;\; Y \in \mathbb{K}^n$$

$$\text{with} \;\; Y \;=\; (y_1, \ldots, y_n)^T \;=\; \Big(\langle y, e_k \rangle\Big)^T_{1 \le k \le n}\,.$$

So to bear in mind the applications of quadratic forms to problems of Euclidean Geometry or Newtonian Mechanics which require the changes of indeterminates to result from changes of Cartesian axes, one could focus the attention

on symmetric matrices and orthogonal transformations P. Under such an appropriate transformation P, there is then a better fact that a quadratic form Q has an associated diagonal matrix, and so has a diagonal form.

In infinite dimensional vector spaces, we still have a duality between quadratic forms and symmetric bilinear forms, but their descriptions are no longer simple and there is no general intrinsec expression for bilinear forms. However, using the theory of Functional Analysis on Hilbert spaces [17],[18], [19],[20],[21],[24],[6], citeyk, we can see that every bounded real bilinear form f defined on an infinite Hilbert space H, admits a bounded linear operator B on H such that

$$f(x,y) \;=\; \langle x, By\rangle\,, \qquad x,\, y \,\in H\,,$$

and so the quadratic form Q defined on H by $Q(x) = f(x,x)$ is also associated with a unique bounded symmetric operator A $\big($equal to $(B+B^*)/2\,\big)$ such that

$$Q(x) \;=\; \langle x, Ax\rangle\,, \qquad x \in H\,.$$

If moreover H is separable, f is symmetric and the operator A associated to f is compact, then H has a countable hilbertian basis $(e_k)_{k\in\mathbb{N}}$ of eigenvectors $(e_k)_{k\in\mathbb{N}}$ of A with respective eigenvalues $(\lambda_k)_{k\in\mathbb{N}}$ converging to zero and such that

$$Ay \;=\; \sum_{k=1}^{+\infty} \lambda_k \langle y, e_k\rangle e_k \qquad \text{for all} \;\; y \in H\,.$$

Therefore the associated quadratic form is expressed by :

$$Q(x) \;=\; \langle x, Ax\rangle \;=\; \sum_{k=1}^{+\infty} \lambda_k |\langle x, e_k\rangle|^2 \qquad \text{for all} \;\; x \in H\,.$$

Note that this class of quadratic form on separable Hilbert spaces is restrictive because given any countable orthonormal system of vectors $(v_k)_{k\in\mathbb{N}}$ of a real Hilbert space H, and a bounded sequence $(\mu_k)_{k\in\mathbb{N}}$ of real numbers (not necessarily convergent), then the functional defined on H by

$$q(x) \;=\; \sum_{k=1}^{+\infty} \mu_k |\langle x, v_k\rangle|^2 \qquad \text{for all} \;\; x \in H\,,$$

is a quadratic form satisfying

$$|q(x)| \;\leq\; \left(\sup_{k\geq 1} |\mu_k|\right) \, ||x||^2\,, \qquad \text{for all} \;\; x \in H\,.$$

The study of quadratic forms can be motivated by the identification of a turning (stationary) point x_0 of a smooth real-valued function f of several variables that leads to the investigation of the value classes of the Hessian

matrix $H_f(a)$ of f at a. The most important value classes are the value class of positive definite forms (which are positive except at 0, in finite dimension) and the value class of negative definite forms (which are negative except at 0, in finite dimension) for the following reasons [1], [2], [3], [23]: Given a natural number n and a real-valued function f of class $\mathcal{C}^2$ defined on a domain $\Omega \subset \mathbb{R}^n$, (i) if a point $x_0 \in \Omega$ is a local optimal (minimum or maximum), then it is a stationary point with a positive or negative semi-definite Hessian $\mathcal{H}_f(x_0)$, (ii) if a stationary point x_0 has a positive or negative definite Hessian $\mathcal{H}_f(x_0)$, then it is a local minimum or maximum (respectively), and (iii) if Ω is convex and the Hessian matrix of f at every point $x \in \Omega$ is positive semi-definite, then f is convex and any statioanary point of f is a global minimum.
While the value class of a quadratic form can be determined from the signs of the eigenvalues of its associated symmetric matrix

Chapter 1

Preliminaries

Throughout the work, $\mathbb{K}$ will hold either for the field of real numbers $\mathbb{R}$ or for the field of complex numbers $\mathbb{C}$.

This part is devoted to a review of some basic notions from Functional Analysis and to the introduction of bilinear forms.. The ideal spaces on which we shall work are Hilbert spaces (e.g. euclidean spaces) that are particular (and even pratical) Banach spaces.

1.1 Banach spaces

Definition 1.1.1

Let X be a linear space over $\mathbb{K}$. A *norm* on X is any nonnegative real-valued function $||\cdot||$ on X satisfying the following conditions :

i) $\forall\, x \in X, \quad ||x|| = 0 \iff x = 0.$ (nondegeneracy)

ii) $||\lambda x|| = |\lambda|\,||x||\,, \quad \forall\, x \in X$ and $\forall\, \lambda \in \mathbb{K}\,.$ (homogeneity)

iii) $||x+y|| \leq ||x|| + ||y||, \quad \forall\, x,\, y \in X\,.$ (subadditivity)

A linear space X endowed with a norm $||\cdot||$ is called a *normed linear space* and is denoted by $(X, ||\cdot||)$.

A normed linear space $(X, ||\cdot||)$ of which norm is not ambiguous will be simply denoted by X.

Definition 1.1.2

Let X be a $\mathbb{K}$-linear space. Two norms $||\cdot||_1$ and $||\cdot||_2$ on X are said to be *equivalent* if there exist positive constants α and β such that

$$\alpha||x||_1 \;\leq\; ||x||_2 \;\leq \beta||x||_1 \quad \forall\, x \in X\,.$$

Definition 1.1.3

Let $(X, ||\cdot||)$ be a normed linear space. A sequence $(x_n)_n$ of elements of X *converges* in $(X, ||\cdot||)$, if there exists an element $a \in X$ such that

$$\lim_{n\to+\infty} ||x_n - a|| \;=\; 0 \;\text{ in } \mathbb{R}\,.$$

In this case a is unique (due to the triangular inequality property of the norm) and we also say that $(x_n)_n$ converges to a with respect to the norm $||\cdot||$ and then write

$$\lim_{n\to+\infty} x_n \;=\; a\,.$$

Definition 1.1.4

Let $(X, ||\cdot||)$ be a normed linear space.

1. A subset F of X is said to be *closed* if every sequence $(a_n)_n$ of elements of A which converges to some element $x \in X$, has its limit x in A. That is:
$$(a_n)_n \subset A \;\text{ converges in } X \quad\Longrightarrow\quad \lim_{n\to+\infty} a_n \in A\,.$$

2. A subset U of X is said to be *open* if its complement; $U^c = X \setminus U$, is closed.

3. The collection $\Im$ of all open sets of the normed linear space $(X, ||\cdot||)$ defines a *topology* on X. In this topological space $(X, \Im)$, $\Im$ is called the norm topology.

4. The *closure* of a subset A of X is the smallest closed set (of X) that contains A.
The closure of A is in fact the intersection of all closed sets of X containing A.
It is denoted by $\overline{A}$ or $\mathrm{cl}(A)$.
A subset $D \subseteq X$ of which closure is equal to X, is said to be *dense* in X.

5. The *interior* of a subset A of X is the largest open set (of X) which is contained in A.
The interior of A is in fact the union of all open sets contained in A.
It is denoted by $\mathring{A}$ or $\mathrm{int}(A)$.

6. A subset $A \subset X$ is said to be bounded if there exists a positive constant M such that
$$x \in A \quad\Longrightarrow\quad ||x|| < M\,.$$
This means that A is contained in some open ball $B(0, M)$.

Now we define the most important topological concept for computational purposes, namely the concept of separability.

Definition 1.1.5

A normed linear space $(X, ||\cdot||)$ is said to be *separable* if it contains a dense subset D which is at most countable.

Proposition 1.1.6

Let X be a $\mathbb{K}$-linear space and assume that $||\cdot||_1$ and $||\cdot||_2$ on X are two equivalent norms on X. Then a sequence of elements of X converges with respect to $||\cdot||_1$ if and only if it does with respect to $||\cdot||_2$.
Therefore two equivalent norms define the same topology on X.
Vice-versa if two norms define the same topology, then they are equivalent. *However two metrics may define the same topology without being strongly equivalent!*

Definition 1.1.7

Let $(X, ||\cdot||)$ be a normed linear space. A sequence $(x_n)_n$ of elements of X is said to be a *Cauchy sequence* if

$$\lim_{m,n\to+\infty} ||x_n - x_m|| \;=\; 0\,; \qquad \text{that is,}$$

$$\forall\varepsilon > 0,\ \exists\, N_\epsilon \in \mathbb{N}\ :\quad ||x_n - x_m|| \;<\; \varepsilon \quad \text{for all } n \geq N_\epsilon \ \text{ and all } \ m \geq N_\epsilon\,.$$

Proposition 1.1.8.
In a normed linear space,

1) every convergent sequence is a Cauchy sequence,

2) and every Cauchy sequence is bounded.

Definition 1.1.9 (Banach space)

A normed linear space $(X, ||\cdot||)$ in which every Cauchy sequence is convergent is called a *Banach space*.

Definition 1.1.10

Given a normed linear space $(X, ||\cdot||_X)$, any Banach space $(E, ||\cdot||_E)$ such that there exists an isometry $j\ :\ X \to E$ with dense range in E; i.e.,

(i) $||j(x)||_E \;=\; ||x||_X$ for all $x \in X$, and

(ii) $\overline{j(X)} = E$,

is called a *completion* of X.

This means that E is a completion of X if E is a Banach space which contains a dense subset isometric to X.

For instance :

i) The completion of $\mathcal{C}([0,1])$ equipped with the norm $||\cdot||_2$ defined by $||f||_2 = \left[\int_0^1 |f(t)|^2 dt\right]^{\frac{1}{2}}$, is (isometric to) $L^2(]0,1[)$.

ii) The completion of $\mathcal{C}^1([0,1])$ equipped with the norm $||\cdot||_{W^{1,2}}$ defined by $||f||_{W^{1,2}} = \left[\int_0^1 |f(t)|^2 dt + \int_0^1 |f'(t)|^2 dt\right]^{\frac{1}{2}}$, is (isometric to) $H^1(]0,1[)$.

Theorem 1.1.11 (Hausdorff)

Every normed linear space has a completion.

Definition 1.1.12

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be two arbitrary normed linear spaces. A map $f : X \to Y$ is said to be continuous at a point $a \in X$ if for every sequence $(x_n)_n$ of X converging to a with respect to $||\cdot||_X$, the sequence $\big(f(x_n)\big)_n$ converges to $f(a)$ in Y with respect to $||\cdot||_Y$.
f is said to be continuous (on X) if it is continuous at every point of X.
Equivalently, f is continuous if and only if the pre-image of every open set in Y is an open set in X.

Recall that given two $\mathbb{K}$-linear spaces X and Y,

- a map or operator
$$T : X \longrightarrow Y$$
is said to be *linear* if for all $x_1, x_2 \in X$ and for all $\alpha_1, \alpha_2 \in \mathbb{K}$ we have
$$T(\alpha_1 x_1 + \alpha_2 x_2) = \alpha_1 T(x_1) + \alpha_2 T(x_2)\,.$$
Equivalently, $T : X \longrightarrow Y$ is linear if for all $x_1, x_2 \in X$ and for all $\alpha \in \mathbb{K}$, we have
$$T(\alpha x_1 + x_2) = \alpha T(x_1) + T(x_2)\,.$$

- a map or operator
$$T : X \longrightarrow Y$$

is said to be *antilinear* if for all $x_1, x_2 \in X$ and for all $\alpha_1, \alpha_2 \in \mathbb{K}$ we have

$$T(\alpha_1 x_1 + \alpha_2 x_2) \;=\; \overline{\alpha_1} T(x_1) \,+\, \overline{\alpha_2} T(x_2)\,.$$

Equivalently, $T : X \longrightarrow Y$ is linear if for all $x_1, x_2 \in X$ and for all $\alpha \in \mathbb{K}$, we have

$$T(\alpha x_1 + x_2) \;=\; \alpha T(x_1) \,+\, T(x_2)\,.$$

Theorem 1.1.13

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces. Then a linear map $T : X \to Y$ is continuous if and only if T is a *bounded linear map* in the sense that there exists a constant real number $\alpha \geq 0$ such that

$$\|T(x)\|_Y \;\leq\; \alpha\, \|x\|_X \qquad \forall x \in X.$$

Notations 1.1.14

Let X and Y be two given arbitrary normed linear spaces.

The set of all bounded linear maps (i.e. continuous linear maps) from X into Y is a linear space that will be denoted by $\mathcal{B}(X, Y)$. When $X = Y$, we simply write $\mathcal{B}(X)$ instead of $\mathcal{B}(X, X)$.

Given a bounded linear map $T : X \to Y$, we shall set

$$\|T\|_{\mathcal{B}(X,Y)} \;=\; \inf\Big\{\, k : \; \|T(x)\|_Y \leq k\|x\|_X \;\forall\, x \in X \,\Big\}$$

that will be simply written as $\|T\|$ when there is no ambiguity.

We denote by $X^* := \mathcal{B}(X, \mathbb{K})$ the topological *dual* of X; that is, the set of all continuous linear functionals of X.
Elements of X^* are also called *continuous linear forms* or *bounded linear forms.*

Proposition 1.1.15

Let $(X, \|\cdot\|_X)$ be a nontrivial normed linear space and $(Y, \|\cdot\|_Y)$ be an arbitrary normed linear space. Then for every $T \in \mathcal{B}(X, Y)$, we have

$$\|T(x)\|_Y \;\leq\; \|T\|\, \|x\|_X \qquad \forall\, x \in X\,,$$

and

$$\|T\| \;=\; \sup_{\|x\|_X \leq 1} \|T(x)\|_Y \;=\; \sup_{\|x\|_X = 1} \|T(x)\|_Y \;=\; \sup_{\|x\|_X \neq 0} \frac{\|T(x)\|_Y}{\|x\|_X}\,.$$

Theorem 1.1.16

Let $(X, \|\cdot\|_X)$ and $(Y, \|\cdot\|_Y)$ be normed linear spaces. Then

1. $\big(\mathcal{B}(X,Y)\,, \|\cdot\|_{\mathcal{B}(X,Y)}\big)$ is a normed linear space.
2. If moreover $(Y, \|\cdot\|_Y)$ is a Banach space, then $\big(\mathcal{B}(X,Y), \|\cdot\|_{\mathcal{B}(X,Y)}\big)$ is a Banach space.

Corollary 1.1.17

The dual X^* of any normed linear space X is (always) a Banach space.

Remark 1.1.18

Given a normed linear space $(X, \|\cdot\|_X)$, the dual X^* being a normed linear space (in fact a Banach space) has also a dual X^{**} called the *bidual* of X.

Moreover there exists a canonical injection $J : X \hookrightarrow X^{**}$ defined by

$$\begin{array}{rccc} J : & X & \longrightarrow & X^{**} \\ & x & \longmapsto & J(x)\,, \end{array}$$

where $J(x)$ is the continuous form on X^* defined by

$$\langle J(x),\, f\rangle \;:=\; \langle f,\, x\rangle \;:=\; f(x)\;;\;\; \forall\, f \in X^*\,.$$

Definition 1.1.19 (Reflexive space)

A normed linear space $(X, \|\cdot\|_X)$ is *reflexive* if it is a Banach space such that the canonical injection $J : X \hookrightarrow X^{**}$ is surjective.

Theorem 1.1.20 (Hahn-Banach)

Let $(X, \|\,.\,\|)$ be a normed linear space and $E \subset X$ be a linear subspace of X. If $f : E \longrightarrow \mathbb{R}$ is a continuous linear functional, then there exists $F \in X^*$ that extends f such that

$$\|F\|_{X^*} = \|f\|_{E^*}\;.$$

Corollary 1.1.21

For every $x_0 \in X$, there exists $f_0 \in X^*$ such that

$$\|f_0\| = \|x_0\| \quad \text{and} \quad \langle f_0, x_0\rangle = \|x_0\|^2.$$

Theorem 1.1.22 (Uniform Boundedness Principle)

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be two Banach spaces and let $\{T_i\}_{i\in I}$ be a family (not necessarily countable) of continuous linear operators from X into Y. Assume that

$$\sup_{i\in I} ||T_i(x)|| < \infty \quad \forall x \in X.$$

Then

$$\sup_{i\in I} ||T_i||_{\mathcal{B}(X,Y)} < \infty.$$

In other words there exists a constant $c > 0$ such that

$$||T_i(x)|| \leq c||x|| \quad \forall x \in X, \ \forall i \in I.$$

Theorem 1.1.23 (Open Mapping Theorem)

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be two Banach spaces and let T be a continuous linear operator from X into Y that is surjective. Then there exists a constant $c > 0$ such that

$$B_Y(0, c) \subset T(B_X(0, 1)).$$

Corollary 1.1.24 (Banach Isomorphism Theorem)

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be two Banach spaces and let T be a continuous linear operator from X into Y that is bijective. Then the inverse map T^{-1} is also continuous (from Y into X).

Theorem 1.1.25 (Closed Graph Theorem)

Let $(X, ||\cdot||_X)$ and $(Y, ||\cdot||_Y)$ be two Banach spaces and let T be a linear operator from X into Y. Assume that the graph of T, $G(T)$, is closed in $E \times F$. Then T is continuous.

Remark 1.1.26

The converse is obviously true, since the graph of any continuous map (linear or not) is closed.

Definition 1.1.27 (Compactness)

A subset A of a Banach space X is said to be *compact* if every cover of A by open sets of X, has a finite subcover.

Proposition 1.1.28

Every compact set of a Banach space is closed and bounded.
The converse is not true in general.

Theorem 1.1.29 (Bolzano-Weierstrass)

A subset A of a Banach space X is said to be compact if and only if any sequence of A has a subsequence that converges to some point of A.

Theorem 1.1.30 (Heine-Borel)

Given a natural number n, a subset of the euclidean space $\mathbb{R}^n$ is compact if and only if it is closed and bounded.

Heine-Borel Theorem fails in infinite dimensional Banach spaces. In fact we have the following characterization of finite dimensional normed spaces over $\mathbb{R}$.

Theorem 1.1.31 (Riesz)[19]

Let E be a Banach space over the field $\mathbb{K}$. Then the closed unit ball of E is compact if and only if the dimension of E is finite.

Thus Riesz Theorem characterizes the compactness of the closed unit ball of a Banach space E by the finiteness of the dimension of E.

Therefore, we need other types of topologies on infinite dimensional spaces

Definition 1.1.32

Let E be a real Banach space. To each f in E^*, we assign the map

$$\begin{array}{rccl} \phi_f : & E & \longrightarrow & \mathbb{R} \\ & x & \longmapsto & \phi_f(x) = \langle f,\, x\rangle. \end{array}$$

We denote the family of all such maps from E into $\mathbb{R}$ by $\{\phi_f\}_{f\in E^*}$.

The *weak topology* on E (denoted by ω) is the smallest topology on E which makes the maps ϕ_f continuous.

Proposition 1.1.33

Let (E, ω) be a real Banach space endowed with the weak topology ω . Then ω is Hausdorff; that is, for any two different points x_1 and x_2 taken in (E, ω), there exists two respective disjoint weakly open neighbourhoods U_1 and U_2; that is U_1, U_2 belongs to ω and $U_1 \cap U_2 = \emptyset$.

Consequently if a sequence $\{x_n\}_{n\in\mathbb{N}}$ in (E, ω) converges weakly to some x (i.e $x_n \rightharpoonup x$) in (E, ω), then x is unique.

Proposition 1.1.34

A sequence $\{x_n\}_{n\in\mathbb{N}}$ in a real Banach space E converges weakly to some x in E if and only if $f(x_n) \longrightarrow f(x)$ for each $f \in E^*$.

Proposition 1.1.35

Given a finite dimensional normed space E, the norm-topology (strong topology) and the weak topology coincide on E.

Theorem 1.1.36 (Eberlein-Smulyan)

A real Banach space E is reflexive if and only if every (norm) bounded sequence in E has a subsequence which converges weakly to an element of E.

On the dual of a normed space, we can define a weaker topology as follows.

Definition 1.1.37

Let E be a real Banach space. Define for every $x \in E$, the map φ_x defined on E^* by

$$\varphi_x(f) \;=\; f(x)\,, \quad \forall\; f \in E^*\,.$$

Then the *weak** topology on E^* is defined as the smallest topology on E^* for which the maps φ_x; $x \in E$, are continuous.

The *weak* topology* of E (denoted by ω^*) is the

Proposition 1.1.38

Let E be a Banach space. A sequence of bounded linear forms $\{f_n\}_{n\in\mathbb{N}}$ of elements of E^* converges weakly* to some $f \in E^*$ if and only if $f_n(x) \longrightarrow f(x)$ for each $x \in E$.

Proposition 1.1.39

Given a finite dimensional normed space E, the norm-topology (strong topology), the weak topology and the weak* topology coincide on $E^a st$.

Theorem 1.1.40 (Banach-Alaoglu)[19]

For every Banach space E, the closed unit ball is weakly* compact.

Now we consider the spectral properties of bounded linear operators.

Definitions 1.1.41 (Spectra and resolvents)

Let E be a Banach space over $\mathbb{K}$ and T be a bounded linear operator of E; i.e., $T \in \mathcal{B}(E)$. The *spectrum* $\sigma(T)$ of T is defined by

$$\sigma(T) \;=\; \{\lambda \in \mathbb{K} : \; \lambda I - T \quad \text{is not invertible} \quad \text{in } \mathcal{B}(E)\}\,.$$

The *resolvent set* $\rho(T)$ of T is the complementary of $\sigma(T)$ in $\mathbb{K}$; that is,

$$\rho(T) \;=\; \mathbb{K} \setminus \sigma(T).$$

The elements of $\rho(T)$ are called the *regular values* of T.
If $\lambda \in \rho(T)$, then the operator $R_\lambda(T) = (\lambda I - T)^{-1}$ is called the *resolvent operator* of T at λ.

The spectrum is decomposed into the disjoint union of the following three sets:

a) The *point spectrum* of T :

$$\sigma_p(T) \;=\; \Big\{\lambda \in \mathbb{K} \;:\; \mathrm{Ker}(\lambda I - T) \neq \{0\}\Big\}.$$

b) The *continuous spectrum* of T :

$$\sigma_c(T) \;=\; \Big\{\lambda \in \mathbb{K} \;:\; \mathrm{Ker}(\lambda I - T) = \{0\},\; \overline{R(\lambda I - T)} = E \text{ but } R(\lambda I - T) \neq E\Big\}.$$

c) The *residual spectrum* of T :

$$\sigma_r(T) \;=\; \Big\{\lambda \in \mathbb{K} \;:\; \mathrm{Ker}(\lambda I - T) = \{0\},\; \overline{R(\lambda I - T)} \neq E\Big\}.$$

Remark/Definition 1.1.42 [24],[19]

1. Given $T \in \mathcal{B}(E)$, an element of $\sigma_p(T)$ is called an *eigenvalue* of T and a non-zero vector v such that $Tv = \lambda v$ is called an *eigenvector* of T associated to the eigenvalue λ. Eigenvalues and eigenvectors are sometimes called *characteristic values* and *characteristic vectors* respectively.

   More generally if $\lambda \in \sigma_p(T)$, $n \in \mathbb{N}$ and v is a nonzero element of E such that $(\lambda I - T)^n v = 0$, then v is called a *principal vector* associated with the eigenvalue λ.

2. Given an eigenvalue λ of an operator T, the *geometric multiplicity* of λ is by definition the dimension of $\mathrm{Ker}(\lambda I - T)$. And the *algebraic multiplicity* of λ is defined as follows:

   - when E is finite-dimensional and equipped with a basis, it is the multiplicity of λ as a root of the characteristic polynomial of the matrix associated to T,
   - more generally for an arbitrary nontrivivial Banach space E, it is the dimension of the vector subspace

   $$\bigcup_{k=1}^{\infty} \mathrm{Ker}(\lambda I - T)^k \,.$$

   Thus the algebraic multiplicity of an eigenvalue is always greater than or equal to its geometric multiplicity.

The notion of spectrum can be defined for unbounded operators defined in a normed space.

Example 1.1.43 [24],[30]

1. Let n be a natural number and $A : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator with corresponding matrix $\mathbf{A}$. then it is well-known that T has n eigenvalues $\lambda_1, \ldots, \lambda_n$ (counted with their multiplicities) and that the eigenvalues are the roots of the characteristic polynomial $P(\lambda) = \det(\mathbf{A} - \lambda I_n)$ for A. Since $A - \lambda I_n)$ is invertible when λ is not an eigenvalue, it follows that the spectrum $\sigma(A)$ of A is a pure point spectrum, namely

   $$\sigma(A) \;=\; \big\{\lambda_k \,:\, 1 \le k \le n\big\} \;=\; \sigma_p(A)\,,$$

   and that the resolvent set of A is the complex plane except finitely many points, namely

   $$\rho(A) \;=\; \mathbb{C} \setminus \big\{\lambda_k \,:\, 1 \le k \le n\big\}\,.$$

2. Let $E = \mathcal{C}([0,1])$ be equipped with the supremum norm and consider the operator $T \,:\, E \longrightarrow E;\; f \longmapsto Tf$ defined by

   $$[Tf](x) \;=\; \int_0^x f(s)\,ds\,, \qquad \forall x \in [0,\,1]\,.$$

   Then it is not hard to check that T has no eigenvalue, T is injective but not surjective and its range is not dense in E, and moreover $T\lambda I$ is invertible in $\mathcal{B}(E)$. Therefore

   $$\sigma(T) \;=\; \{0\} \;=\; \sigma_r(T)\,.$$

Theorem 1.1.44 [20]

Let E be a Banach space over $\mathbb{C}$ and $T \in \mathcal{B}(E)$. Then the following holds for the spectrum of T.

i) $\sigma(T)$ is a closed subset of $\mathbb{C}$

ii) $\sigma(T) \subset B(0, ||T||_{\mathcal{B}(X)})$

iii) $\sigma(T)$ is a compact subset of $\mathbb{C}$

Definition 1.1.45 (Spectral radius) [20]

Let E be a Banach space over $\mathbb{C}$ and $T \in \mathcal{B}(E)$. Then the *spectral radius* of T is defined by

$$r(T) \;:=\; \sup\Big\{|\lambda| \;:\; \lambda \in \sigma(T)\Big\} \;=\; \max\Big\{|\lambda| \;:\; \lambda \in \sigma(T)\Big\}.$$

Futhermore, we have the Gelfang formula

$$r(T) \;=\; \lim_{n\to\infty} \|T^n\|^{\frac{1}{n}} \;=\; \inf_{n\in\mathbb{N}} \|T^n\|^{\frac{1}{n}} \;.$$

Definition 1.1.46 (Compact linear maps)

Let E and F be two Banach Spaces over $\mathbb{K}$. A linear map, $T : E \to F$ is said to be compact if the image of the closed unit ball $\overline{B}_E(0,1)$ by T is a relatively compact subset of F.
In other words, T is compact if $\overline{T\big(\overline{B}_E(0,1)\big)}$ is compact.
This definition is equivalent to each of the following properties.

i) For each bounded subset $B \subset E$, the image set $T(B)$ is relatively compact in F.

ii) For every bounded sequence $\{x_n\}_{n\in\mathbb{N}} \subset E$, the sequence $\{Tx_n\}_{n\in\mathbb{N}}$ has a convergent subsequence in F.

Proposition 1.1.47 (Properties of compact linear maps) [20]

1. Every compact linear map is bounded.

2. For all Banach spaces E and F over the field $\mathbb{K}$, the set $\mathcal{K}(E,F)$ of compact linear maps from E into F is a linear subspace of the space $\mathcal{B}(E,F)$ of bounded linear maps from E into F.

3. Let E, F and G be Banach spaces, and let $T : E \longrightarrow F$, $S_1 : F \longrightarrow G$ and $S_2 : G \longrightarrow E$ be bounded linear operators. Then

i) If the range of T; $R(T)$, is finite dimensional, then T is compact. Therefore, every linear map defined on a finite dimensional normed space is not only bounded, but also compact.

ii) If T is compact, then $S_1 \circ T$ and $T \circ S_2$ are compact. Therefore the set $\mathcal{K}(E)$ of compact linear operators (endomorphisms) of E is a two-sided ideal of the algebra $\mathcal{B}(E)$ of bounded linear operators (endomorphisms) of E.

iii) The limit of a convergent sequence of compact linear operators with respect to the operator norm, is compact. Thus $\mathcal{K}(E)$ in closed in $\mathcal{B}(E)$.

Next, we state the Riesz-Schauder spectral theory of compact linear operators.

Theorem 1.1.48 (Riesz-Schauder spectral theory)[19],[22]

Let E be a complex Banach space and T be a compact linear operator of E. Then:

1. The spectrum of T consists of an at most countable set of points of the complex plane which has no point of accumulation except possibly $\lambda = 0$.

2. Every nonzero number of the spectrum of T is an eigenvalue of T of finite multiplicity.

3. The dual operator of T denoted by $T^* \in \mathcal{B}(E^*)$ and defined by

$$T^*(f) \;=\; f \circ T \quad \text{for every } f \in E^*\,,$$

is also compact and a nonzero number is an eigenvalue of T if and only if it is an eigenvalue of T^*.

Remark 1.1.49 [22]

The notion of dual operator is an extension of the notion of transposed matrix. In fact, if E is finite dimensional and equipped with a given basis, then the matrix of the dual of a linear operator of E is the transposed matrix of the matrix of T. And More generally, if E is a normed space and $T \in \mathcal{B}(E)$, then using the duality pairing of $E^* \times E$, we have

$$\langle f,\, Tx\rangle \;=\; \langle T^* \circ f,\, x\rangle \quad \forall\, x \in E\,.$$

The above Theorem 1.1.48 can be rephrased as follows.

Theorem 1.1.50

Let E be a complex Banach space and T be a compact linear operator of E.

1. If $\lambda \in \sigma(T)$ and $\lambda \neq 0$, then λ is an eigenvalue of T of finite multiplicity.
2. $\sigma(T)$ is either finite or countably infinite.
3. If $\sigma(T)$ is infinite, then

$$\sigma(T) \;=\; \{0\} \cup \big\{\lambda_n \;:\; n = 1,\, 2,\, \dots\big\}\,;$$

   where $\{\lambda_n\}_{n\in\mathbb{N}}$ is a sequence of complex numbers converging to 0.

As a corollary we have

Theorem 1.1.51 [20]

Let E be an infinite dimensional complex Banach space and $T \in \mathcal{B}(E)$ be compact. Then The following holds.

1. $0 \in \sigma(T)$
2. $\sigma(T) \setminus \{0\}$ consists of eigenvalues of finite multiplicity.
3. $\sigma(T) \setminus \{0\}$ is either empty, finite or a sequence of complex numbers converging to 0. That is $\sigma(T) \setminus \{0\}$ is a discrete set with no limit point other than 0.

1.2 Hilbert spaces

Definition 1.2.1 (Inner product)

An *inner product* on a $\mathbb{K}$-linear space E is any functional $\langle\,,\,\rangle$ defined on $E \times E$ which is a positive hermitian and nondegenerate form; that is,

$$\begin{array}{rccc} \langle\,,\,\rangle : & E \times E & \longrightarrow & \mathbb{K} \\ & (x, y) & \mapsto & \langle x,\, y\rangle \end{array}$$

and satisfies the following conditions :

1. $\langle x,\, x\rangle \;\geq\; 0 \;\; \forall\, x \in E,$ and $\langle x,\, x\rangle \;=\; 0$ if and only if $x = 0$.
2. $\langle y,\, x\rangle \;=\; \overline{\langle x,\, y\rangle} \quad \forall\, x,\, y \in E.$
3. $\langle \alpha_1 x_1 + \alpha_2 x_2,\, y\rangle \;=\; \alpha_1\langle x_1,\, y\rangle \;+\; \alpha_2\langle x_2,\, y\rangle \qquad \forall\, x_1,\, x_2,\, y \in E$ and $\forall\, x_1,\, \alpha_2 \in \mathbb{K}$.

Remark 1.2.2

1. A form $\Phi : E \times E \to \mathbb{K}$ which is linear with respect to its first argument and antilinear with respect to its second argument is said to be *sesquilinear*.
   If $\Phi : E \times E \to \mathbb{K}$ is sesquilinear and satisfy
   $$\Phi(x, y) = \overline{\Phi(y, x)} \quad \forall\, x, y \in E\,,$$
   then it is also called a *hermitian sesquilinear* form.

2. When $\mathbb{K} = \mathbb{R}$; that is, E is a real vector space, a hermitian (sesquilinear) form on E is just called a *symmetric bilinear form* on E.

Definition 1.2.3

A linear space endowed with an inner product is called an *inner product space* or a *prehilbertian* space.

A finite dimensional real prehilbertian space is also called a *euclidean space*

Theorem 1.2.4 (Cauchy-Schwarz-Bunyakovsky Inequality)

Let $(E, \langle\, ,\, \rangle)$ be an inner product space. Then

$$|\langle x, y\rangle|^2 \leq \langle x, x\rangle \langle y, y\rangle \qquad \forall\, x, y \in E.$$

The equality of this inequality holds if and only if x and y are linearly dependent.

Theorem 1.2.5 (Norm induced by an inner product)

Let $(E, \langle\, ,\, \rangle)$ be an inner product space. Then the function

$$\begin{array}{rcccl} ||\, x\, ||_E & : & E & \longrightarrow & \mathbb{R} \\ & & x & \longmapsto & \sqrt{\langle x, x\rangle} \end{array}$$

defines a norm on E.

Definition 1.2.6 (Hilbert space)

An inner product space E is called a Hilbert space (usually denoted by H), if $(E, ||\,.\,||_E)$ is a Banach space.

Remark 1.2.7 (Hilbert space)

Finite dimensional real Hilbert spaces are also called Euclidean spaces. And since all finite dimensional normed spaces are Banach spaces, there is no difference between finite dimensional real prehilbertian spaces and finite dimensional real Hilbert spaces.

Theorem 1.2.8 (Parallelogram law and Polarization identity)

Let $(E, \langle\, , \rangle)$ be an inner product space. Then we have

1. the parallelogram law :
$$||x+y||^2 + ||x-y||^2 \;=\; 2(||x||^2 + ||y||^2) \qquad \forall\, x, y \in E\,,$$
2. and the Polarization identity :
$$\langle x, y\rangle \;=\; \frac{1}{4}\Big(||x+y||^2 - ||x-y||^2 + i||x+iy||^2 - i||x-iy||^2\Big) \qquad \forall\, x, y \in E\,.$$

Note that the parallelogram law characterizes the norms that are induced by inner products according to a theorem by Von Neumann.

Definitions 1.2.9 (Orthogonality)

Let $(E, \langle\, , \rangle)$ be an inner product space.
Two vectors x and y in E are said to be orthogonal (written $x \perp y$ and read x 'perp' y) if $\langle x, y\rangle \;=\; 0$.

If M is a non empty subset of E, we write $x \perp M$ (and read x orthogonal to M) if x is orthogonal to every element of M.

Given a non-empty subset M of E, we denote by $M^\perp$, the set of all elements of E which are orthogonal to M. That is,
$$M^\perp \;=\; \Big\{x \in E\,,\;\; \langle x, y\rangle = 0 \quad \forall\, y \in M\Big\}\,.$$
The set $M^\perp$ is then called the *orthogonal* of M, it is a closed vector subspace of E.

Theorem 1.2.10 (Projection Theorem)

Let H be a Hilbert space and M a closed subspace of H. For arbitrary vector x in H, there exists a unique vector $y^\star \in M$ such that,
$$||x - y^\star||_H \;=\; \inf_{y \in M} ||x - y||_H.$$
Furthermore, $\overline{x} \in M$ is such a vector if and only if $(x - \overline{x}) \perp M$.
The Projection Theorem yields the following definition end theorems.

Definition 1.2.11 (Direct Sum of vectors spaces)

Let X and Y be two subspaces of a vector space E. Then E is said to be the *direct sum* of X and Y if

$$E = X + Y \text{ and } X \cap Y = \{0\} \ .$$

This means that every vector $u \in E$ has a unique decomposition of the form $u = x + y$ with $x \in X$ and $y \in Y$.
in this case we write $E = X \oplus Y$.

Theorem 1.2.12 (Direct Sum Decomposition)

Let F be a closed subspace of a Hilbert space H. Then,

$$H \ = \ F \oplus F^{\perp} \, .$$

Therefore we say that $F^{\perp}$ is the orthogonal complement of F

Theorem 1.2.13 (Riesz Representation)

Let H be a Hilbert space and let f be a bounded linear functional on H. Then,

(i) There exists a unique vector y_o in H such that

$$f(x) \ = \ \langle x, y_o \rangle, \qquad \text{for every } \ x \in H.$$

(ii) Moreover, $||f||_{H^*} = ||y_o||_H$.

Corollary 1.2.14

If H be a Hilbert space, then $H \ \simeq \ H^*$ via the canonical map

$$\begin{array}{rccc} \varphi : & H & \longrightarrow & H^* \\ & a & \longmapsto & \varphi_a \, , \end{array}$$

where φ_a is defined by

$$\varphi_a \ = \ \langle x, \, a \rangle \, , \qquad \forall \, x \in H \, .$$

This canonical map is bijective, isometric and antilinear.

Corollary 1.2.15

Real Euclidean and Hilbert spaces are reflexive.

Definitions 1.2.16 (Hilbertian basis)

Let H be a Hilbert space.

1. A *unit vector* of H is a vector of H is equal to 1.

2. A family $\{u_\alpha\}_{\alpha\in\Gamma}$ of nonzero vectors of H is said to be *orthogonal*, if the vectors of the family are pairwise orthogonal; that is,
$$\langle u_\alpha\,,\,u_\beta\rangle \;=\; 0\,,\quad \text{for all}\;\; \alpha\neq\beta \text{ in } \Gamma\,.$$

3. A family $\{u_\alpha\}_{\alpha\in\Gamma}$ of (nonzero) vectors of H is said to be *orthonormal* if these vectors are all unit vectors and pairwise orthogonal; that is,
$$\begin{cases} ||u_\alpha|| \;=\; \sqrt{\langle u_\alpha\,,\,u_\alpha\rangle} \;=\; 1\,, & \text{for all}\;\; \alpha\in\Gamma. \\ \langle u_\alpha\,,\,u_\beta\rangle \;=\; 0\,, & \text{for all } \alpha\neq\beta \text{ in } \Gamma\,. \end{cases}$$

4. A family $\mathscr{A}$ of vectors of H is said to be *total* or *complete* if the vector subspace spanned by $\mathscr{A}$ is dense is H; that is,
$$\overline{\mathrm{Span}(\mathscr{A})} \;=\; H\,.$$

5. A family of vectors of H that is both orthonormal and complete is called a *hilbertian basis*.
A hilbertian basis can be finite, countable or uncountable.

   The crucial difference between a hilbertian basis and a (finite) orthonormal basis is that in the first case the expansion of a vector may not be a linear combination of some of the elements of the hilbertian basis, but a series of vectors!

Examples 1.2.17

In the real space
$$\ell^2 \;=\; \left\{ u=(u_n)_{n\geq 1}\subset\mathbb{R}\,;\quad \sum_{n=1}^{\infty} u_n^2<\infty \right\}$$
endowed with the inner product defined by
$$\langle u,\,v\rangle \;=\; \sum_{n=1}^{\infty} u_n v_n\,,$$
the following vectors orthonormal and complete.
$$\begin{aligned} e_1 \;&=\; (1,0,0,0,\ldots) \\ &\vdots \\ e_k \;&=\; (\delta_{k,n})_n \quad \text{where}\;\; \delta_{k,n}=1 \text{ si } n=k \;\;\text{and}\;\; \delta_{k,n}=0 \text{ otherwise} \\ &\vdots \end{aligned}$$

Proposition 1.2.18

Let H be a Hilbert space. Then

1. For every vector subspace F of H, we have

$$\left(F^{\perp}\right)^{\perp} \;=\; \overline{F}\,.$$

   In particular if F is a closed subspace, then $\left(F^{\perp}\right)^{\perp} \;=\; F\,.$

2. For every nonempty subset A of H, $A^{\perp}$ is a closed vector subspace of H (and so a Hilbert subspace) and

$$\left(A^{\perp}\right)^{\perp} \;=\; \overline{\mathrm{Span}(A)}\,.$$

3. A nonempty set $\mathcal{S}$ of H is complete (or total) if and only if

$$\mathcal{S}^{\perp} \;=\; \{0\}\,.$$

Theorem 1.2.19 [17]

Let $(H,\ \langle\ ,\ \rangle)$ be a Hilbert space. Then

1. $\mathcal{H}$ has a hilbertian basis (i.e., an *orthonormal basis*) that can be finite countable or uncountable. Furthermore, every orthonormal set in H is contained in some hermitian basis.

2. A hilbertian basis of H is at most countable if and only if H is separable; that is, H contains a dense subset that is at most countable.

3. For every orthonormal sequence of vectors $\{e_k\}_{k\in\mathbb{N}}$ of elements of H, we have *Bessel Inequality*:

$$\forall\, x \;\in\; H, \qquad \sum_{k=1}^{+\infty} |\langle x, e_k\rangle|^2 \;\leq\; ||x||^2\ .$$

4. If H is separable and $\{e_k\}_{k\in\mathbb{N}^*}$ is a <u>complete and orthonormal</u> sequence of vectors of H, then we have *Parseval Identity*:

$$\forall\, x \;\in H, \qquad \sum_{k=1}^{+\infty} |\langle x, e_k\rangle|^2 \;=\; ||x||^2\ .$$

Proposition 1.2.20

Let $(H,\ \langle\,,\,\rangle)$ be an infinite dimensional Hilbert space and $\{e_k\}_{k\in\mathbb{N}}$ be an orthonormal set of H. Then

1. For every $x \in H$, we have
$$\lim_{k\to+\infty} \langle x,\, e_k\rangle \;=\; 0\,.$$

2. Every bounded sequence of scalars $\{\mu_k\}_{k\in\mathbb{N}}$ gives a bounded linear operator T on H defined by
$$Tx \;=\; \sum_{k=1}^{+\infty} \mu_k \langle x,\, e_k\rangle e_k\,, \quad \forall x \in H\,,$$
of which norm is $||T|| = \sup_{k\geq 1} |\mu_k|$.

Proof

1. Follows from the convergence of the numerical series $\sum_{k=1}^{+\infty} |\langle x,\, e_k\rangle|^2$ according to Bessel inequality.

2. Follows from the convergence in H of the series $\sum_{k=1}^{+\infty} \mu_k \langle x,\, e_k\rangle|^2$ and the bounds
$$||Tx|| \;\leq\; \sup_{k\geq 1} |\mu_k|\, ||x||\,, \quad \forall x \in H$$
and
$$\sup_{k\geq 1} ||Te_k|| \;=\; \sup_{k\geq 1} |\mu_k|\,.$$

Theorem 1.2.21 (Construction of compact linear operators) [24]

Let $(H,\ \langle\,,\,\rangle)$ be an infinite dimensional, separable complex Hilbert space with orthonormal basis $(e_k)_{k\in\mathbb{N}}$ and let $(\lambda_k)_{k\in\mathbb{N}}$ be an arbitrary sequence of complex numbers. For every $x \in H$, consider the series
$$Tx \;=\; \sum_{k=1}^{+\infty} \lambda_k \langle x,\, e_k\rangle e_k\,.$$

Then

1. The series is convergent and the sum defines a linear operator T on H if the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is bounded.

2. T exists and is bounded if and only if the sequence $(\lambda_k)_{k\in\mathbb{N}}$ is bounded.

   When $\lambda_k \in \{0,\, 1\}$ for all k, T is an orthogonal projection.

3. T exists and is a compact linear operator if and only if the sequence $(\lambda_k)_{k\in\mathbb{N}}$ converges to 0.

   When only finitely many of the terms λ_k are nonzero, T is a finite rank operator.

Definition 1.2.22 (Numerical range)[19]

Let H be a complex Hilbert space and T be a bounded linear operator on H. The *numerical range* of T is defined by

$$W(T) \;=\; \Big\{ \langle Tx,\, x\rangle \;:\;\; x \in H,\; ||x|| = 1 \Big\}.$$

Proposition 1.2.23 [19]

Let H be a complex Hilbert space and T be a bounded linear operator on H. Then

$$\sigma(T) \;\subset\; \overline{W(T)}\,;$$

where

$$W(T) \;=\; \Big\{ \langle Tx,\, x\rangle \;:\;\; x \in H,\; ||x|| = 1 \Big\}..$$

More precisely, if $\lambda \notin \overline{W(T)}$, then $\lambda \in \rho(T)$ with

$$\big\|(T - \lambda I)^{-1}\big\| \;\leq\; \frac{1}{\operatorname{dist}(\lambda,\, W(T)}\,.$$

Definition 1.2.24 [6],[27]

Let H be a complex Hilbert space and T be a bounded linear operator on H. The *numerical radius* of T is defined by

$$w(T) \;=\; \sup\Big\{ \,|\langle Tx,\, x\rangle| \;:\;\; x \in H,\; ||x|| = 1 \Big\}.$$

It satisfies the following inequalities

$$r(T) \;\leq\; w(T) \quad \text{and} \quad \frac{||T||}{2} \;\leq\; w(T) \;\leq ||T||\,.$$

As a result of the Riesz representation Theorem, we have the following characterization of weak convergence in a Hilbert space.

Proposition 1.2.25

Let H be a Hilbert space.
A sequence $\{x_n\}_{n\in\mathbb{N}}$ of elements of H converges weakly to some $a \in H$ if and only if

$$\lim_{n\to+\infty} \langle x_n,\, y\rangle \;=\; \langle a,\, y\rangle\,, \qquad \forall\, y \in H\,.$$

Theorem 1.2.26 [24],[6]

Let H be a Hilbert space and $(x_n)_{n\in\mathbb{N}}$ be a weakly convergent sequence with limit point x. then for every compact linear operator T on H, the image sequence $\big(Tx_n\big)_{n\in\mathbb{N}}$ converges strongly (in norm) to Tx.
That is, for any compact linear operator $T \in \mathcal{B}(H)$, we have

$$x_n \rightharpoonup x \text{ for } n \to +\infty \qquad \Longrightarrow \quad Tx_n \to Tx \text{ for } n \to +\infty\,.$$

Corollary 1.2.27

Let $(e_n)_{n\in\mathbb{N}}$ be an orthonormal sequence of a Hilbert space H. Then for every compact linear operator T of H, we have

$$\lim_{n\to+\infty} Te_n \;=\; 0 \text{ with respect to the norm topology of } H\,.$$

Proof. By Bessel inequality, we have that $(e_n)_{n\in\mathbb{N}}$ converges weakly to 0 in H. Therefore the corollary follows from the above Theorem [ ]. □

1.3 Self-adjoint operators

Let n be a natural number and $T : \mathbb{C}^n \to \mathbb{C}^n$ be a linear operator. Then T can be represented by a complex square matrix of order n. Suppose moreover that T is self-adjoint. Then the matrix associated to T is conjugate symmetric and it is well-known from basic Linear Algebra that all the eigenvalues of this matrix are all real and that there exists an orthonormal basis $(e_1, \ldots, e_n)$ for $\mathbb{C}^n$ in which the matrix of T is diagonal, meaning also that the linear operator T can be expressed as follows :

$$Tx \;=\; \sum_{k=1}^{n} \lambda_k \langle x,\, e_k\rangle e_k\,, \qquad \forall\, x \in \mathbb{C}^n\ ,$$

where $\langle x,\, e_k\rangle$ coincide with k^{th} coordinate of x in the orthonormal basis $(e_1, \ldots, e_n)$ of eigenvectors for T corresponding respectively to the eigenvalues $\big\{\lambda_k\big\}_{1\le k\le n}$.

For infinite dimensional Hilbert spaces, the situation is much more tremendous, but for self-adjoint compact linear operators, a corresponding theory can be well developped. It will be culminated in the famous spectral theorem.

Definition 1.3.1 (Symmetric or self-adjoint operators) [19],[18],[20],[24]

Let H be a Hilbert space over $\mathbb{K}$ and $A : H \to H$ be a bounded linear operator. A is said to be *symmetric* or *self-adjoint* if

$$\langle Ax,\, y\rangle \;=\; \langle x,\, Ay\rangle\ , \qquad \forall\, x,\, y \in H\,.$$

Remark 1.3.2 [24],[22]

1. Given an arbitrary bounded linear operator T on a Hilbert space H, the *adjoint operator* of T is the unique bounded linear operator of H denoted by T^* and satisfying

$$\langle Tx,\, y\rangle \;=\; \langle x,\, T^*y\rangle\;, \qquad \forall\, x,\, y \in H\,.$$

   The existence and properties of T^* are based on the Riesz representation theorem and can be proved following the idea of the proof of [Representation theorem of a real bounded bilinear form].
   Therefore a bounded linear operator T on a Hilbert space is symmetric or self-adjoint if $T = T^*$.

2. Let H be a Hilbert space. Then any map T defined from H into H that satisfies

$$\langle Tx,\, y\rangle \;=\; \langle x,\, T^*y\rangle\;, \qquad \forall\, x,\, y \in H\,,$$

   is necessarily linear and bounded. In fact :

   a) the linearity holds since for all $x,\, y \in H$ and for all $\alpha,\, \beta \in \mathbb{K}$, we have

$$\Big(T(\alpha x + \beta y) \;-\; \alpha Tx \;-\; \beta Ty\Big) \;\in\; H^{\perp} \;=\; \{0\}\,,$$

   b) and the boundedness follows from the Closed graph Theorem [...]

3. Given a Hilbert space H and a linear operator T defined from a dense domain $D(T) \subset H$ into H, the *adjoint operator* of T is defined as the unique operator T^* with domain $D(T^*)$ and such that

$$\langle Tx,\, y\rangle \;=\; \langle x,\, T^*y\rangle\;, \qquad \forall\, x \in D(T) \text{ and } \forall y \in D(T^*)\,.$$

   In this case, a linear unbounded operator T defined on a dense domain $D(T)$ of a Hilbert space H is said to be *symmetric* if T^* is an extension of T, this amounts to:

$$\langle Tx,\, y\rangle \;=\; \langle x,\, T^*y\rangle\;, \qquad \forall\, x,\, y \in D(T)\,.$$

   And again, a linear unbounded operator T defined on a dense domain $D(T)$ of a Hilbert space H is said to be *self-adjoint* if $T^* = T$. This means that not only T is symmetric, but also $D(T) = D(T^*)$.

Proposition 1.3.3 [24],[19]

The numerical range of any bounded, self-adjoint linear operator on a complex Hilbert space is a subset of $\mathbb{R}$.

In particular for any bounded, self-adjoint linear operator T on a Hilbert space over $\mathbb{K}$, $\langle Tx,\, x\rangle$ is a real number for all $x \in H$.

Proof. For all $x \in H$, we have

$$\langle Tx,\, x\rangle \;=\; \langle x,\, Tx\rangle \;=\; \overline{\langle Tx,\, x\rangle}\,.$$

□

Proposition 1.3.4 [24]

Let T be a bounded self-adjoint operator on a complex Hilbert space H. Then all its eigenvalues are real numbers. Furthermore, any pair of eigenvectors corresponding to different eigenvalues are orthogonal.

Proof. This is a well-known result.

- If $Tv = \lambda v$ for $\lambda \in \mathbb{C}$ and $v \in H \setminus \{0\}$, we get

  $$\lambda\langle v,\, v\rangle \;=\; \langle \lambda v,\, v\rangle \;=\; \langle Tv,\, v\rangle \;=\; \langle v,\, Tv\rangle \;=\; \langle v,\, \lambda v\rangle \;=\; \overline{\lambda}\langle v,\, v\rangle\,.$$

  And since $\langle v,\, v\rangle = ||v||^2 \neq 0$, we get $\lambda = \overline{\lambda}$ meaning that λ is real. Note that roughly prouving, λ is real because $\langle Tv,\, v\rangle$ is real and $\langle v,\, v\rangle$ is a nonzero real number.

- If $\lambda_1 \neq \lambda_2$ are two different eigenvalues corresponding respectively to two eigenvectors v_1 and v_2, then λ_1 and λ_2 are real numbers and we have

  $$\begin{aligned} \lambda_1\langle v_1,\, v_2\rangle \;=\; \langle \lambda_1 v_1,\, v_2\rangle \;&=\; \langle Tv_1,\, v_2\rangle \\ &=\; \langle v_1,\, Tv_2\rangle \\ &=\; \langle v_1,\, \lambda_2 v_2\rangle \;=\; \overline{\lambda_2}\langle v_1,\, v_2\rangle \;=\; \lambda_2\langle v_1,\, v_2\rangle\,. \end{aligned}$$

  Therefore

  $$(\lambda_1 \,-\, \lambda_2)\langle v_1,\, v_2\rangle \;=\; 0$$

  and so $\langle v_1,\, v_2\rangle \;=\; 0$ since $\lambda_1 - \lambda_2 \neq 0$.

□

We have the following alternative formular for the norm of a bounded sefl-adjoint operator.

Proposition 1.3.5 [24],[19]

Let T be a bounded self-adjoint operator on a Hilbert space H (over $\mathbb{K}$). Then

$$||T|| \;=\; \sup_{x\in H,\, ||x||=1} |\langle Tx,\, x\rangle|\,.$$

Proof.(Sketch)
Set $\mu = \sup_{||x||=1} |\langle Tx,\, x\rangle|$.

- First of all we have clearly $\mu \leq ||T||$.
- Let $x,\, y \in H$. By expending $\langle T(x+y),\, x+y\rangle$ and $\langle T(x-y),\, x-y\rangle$ we get

$$\begin{aligned}\langle Tx,\, y\rangle \,+\, \langle Ty,\, x\rangle \;&=\; \tfrac{1}{2}\Big(\langle T(x+y),\, x+y\rangle \,-\, \langle T(x-y),\, x-y\rangle\Big)\\ &\leq\; \tfrac{\mu}{2}\left(||x+y||^2 \,+\, ||x-y||^2\right)\\ &\leq\; \mu\left(||x||^2 \,+\, ||y||^2\right) \quad \text{(using parallelogram law).}\end{aligned}$$

  And so

$$\langle Tx,\, y\rangle \,+\, \langle Ty,\, x\rangle \;\leq\; \mu\left(||x||^2 \,+\, ||y||^2\right)\,.$$

- If $x \in H$ is such that $Tx \neq 0$, then by setting $y = \frac{||x||}{||Tx||}Tx$ and by applying the inequality of the previous step, we get

$$||Tx|| \;\leq\; \mu||x||\,.$$

  The latter inequality is also (directly) satisfied even tough $Tx = 0$. Therefore,

$$||Tx|| \;\leq\; \mu||x||\,, \quad \forall x \in H\,,$$

  yielding $||T|| \leq \mu$.

□

Proposition 1.3.6 [18]

Let T be a bounded self-adjoint operator on a real Hilbert space H. Define the lower and upper bounds

$$m \;=\; \inf_{x\in H,\, ||x||=1} \langle Tx,\, x\rangle \qquad \text{and} \quad M \;=\; \sup_{x\in H,\, ||x||=1} \langle Tx,\, x\rangle\,.$$

Then

1. $\sigma(T) \subset [m,\, M]$.
2. $m,\, M \in \sigma(T)$.
3. $||T|| = \max\{-m,\, M\}$.

Corollary 1.3.7 [18]

1. For every bounded linear self-adjoint operator T on a real Hilbert space, either $||T||$ or $-||T||$ is an approximate eigenvalue (i.e., a limit of a sequence of eigenvalues of T).
2. Consequently, every nonzero compact linear self-adjoint operator T on a real Hilbert space, has either $||T||$ or $-||T||$ as an eigenvalue; that is,

$$\big\{-||T||,\, ||T||\big\} \cap \sigma_p(T) \;\neq\; \emptyset\,.$$

1.4 Spectral decomposition

This section deals with the spectral decomposition of a compact linear self-adjoint operator on a separable Hilbert space.

Theorem 1.4.1 (Hilbert-Schmidt) [24], [18], [28]

Let T be a compact self-adjoint operator on a separable Hilbert space H of finite or infinite dimension. Then H admits an at most countable orthonormal basis consisting of eigenvectors for T. More precisely

1. In the finite dimensional case, the numbering of the finite sequence of basis vectors $(e_1, ..., e_n)$ can be chosen such that the corresponding finite sequence of eigenvalues $(\lambda_1, ..., \lambda_n)$ decreases numerically (in absolute value) :
$$|\lambda_1| \;\geq\; |\lambda_2| \;\geq\; \ldots \;\geq\; |\lambda_n|\,.$$
And in the basis of eigenvectors, the operator T is described by :
$$Tx \;=\; \sum_{k=1}^{n} \lambda_k \langle x,\, e_k\rangle e_k \qquad \text{for all} \;\; x \in H = \mathrm{Span}\big\{e_k\,,\; k = 1, ..., n\big\}\,.$$

2. In the separable, infinite dimensional case :

   a) If $T = 0$, then any orthonormal basis of the separable Hilbert space H is a countable orthonormal basis consisting of eigenvectors of T.

   b) If $T \neq 0$, is of finite rank $n \geq 1$, then
   $$H \;=\; \ker(T) \oplus R(T) \quad \text{with} \;\; [\ker(T)]^{\perp} = R(T) \text{ and } \dim[R(T)] = n\,.$$
   In this case $R(T)$ is finite dimensional and invariant under T. Therefore $R(T)$ has a finite orthonomal basis $(e_1, ..., e_n)$ consisting of eigenvectors of the restriction of T to its range $R(T)$ such that their corresponding eigenvalues satisfy
   $$|\lambda_1| \;\geq\; |\lambda_2| \;\geq\; \ldots \;\geq\; |\lambda_n| \qquad \text{(all nonzero).}$$
   Again T is described by
   $$Tx \;=\; \sum_{k=1}^{n} \lambda_k \langle x,\, e_k\rangle e_k \qquad \text{for all} \;\; x \in H\,.$$
   By adding to $(e_1, ..., e_n)$ an orthonormal basis of $\ker(T)$ (also separable), we shall obtain a countable orthonormal basis for H consisting of eigenvectors of T.

c) Otherwise $\big(R(T)$ is infinite dimensional$\big)$, we can find an infinite sequence of orthonormal eigenvectors $\big(e_k\big)_{k\in\mathbb{N}}$; in fact an orthonormal basis of $\overline{R(T)}$ consisting of eigenvectors, with a corresponding sequence of nonzero eigenvalues $\big(\lambda_k\big)_{k\in\mathbb{N}}$ that decreases numerically and tends to 0;

$$|\lambda_1| \;\geq\; |\lambda_2| \;\geq\; \ldots\; |\lambda_k| \;\geq\; \ldots \qquad \text{with} \quad \lim_{k\to+\infty} \lambda_k = 0\,,$$

and for which T can be described as

$$Tx \;=\; \sum_{k=1}^{+\infty} \lambda_k \langle x,\, e_k\rangle e_k \qquad \text{for all} \;\; x \in H\,.$$

Note in this case that

$$\sigma_p(T) \subset \{0\} \cup \big\{\lambda_k \;:\; k = 1,\, 2,\, \ldots\,\big\}\,.$$

Proof (Sketch).
If the range of T is finite dimensional, then we can proceed by induction.
Now we only consider the case in which the range of T is infinite dimensional.

- Set $\mu_0 = 0$ and let
$$\big\{\mu_1\,,\; \mu_2\,,\; \ldots\,\big\}$$
be the countable set of all the nonzero eigenvalues of T. Consider the subspace
$$H_0 \;=\; \ker(T) \qquad \text{(that may be null)}\,,$$
and the eigenspaces
$$H_k \;=\; \ker(T - \mu_k I)\,; \quad k = 1,\, 2,\, \ldots\,.$$
For $k \geq 1$, $\dim(H_k) \geq 1$ since H_k is an eigenspace, and $\dim(H_k) < +\infty$ since T is compact (cf. Theorem 1.1.48, by which every nonzero eigenvalue μ_k must have a finite multiplicity). Thus
$$1 \;\leq\; \dim(H_k) \;<\; +\infty\,, \qquad \forall\, k \geq 1\,.$$

1. For any $m \neq n$ in $\big\{0,\, 1,\, 2\;\ldots\big\}$, H_m and H_n are orthogonal. (See the proof of theorem ...)

2. We have $H \;=\; \overline{\oplus_{k\geq 0} H_k}$ where (recall)
$$\oplus_{k\geq 0} H_k \;=\; \big\{v_{k_0} + \ldots + v_{k_n} \;:\; n \geq 0,\; 0 \leq j \leq n,\; k_j \geq 0,\; v_{k_j} \in H_{k_j}\big\}\,.$$
To see this, set
$$V \;=\; \overline{\oplus_{k=0} H_k}$$

and suppose by contradiction that $V \neq H$. Therefore $V^\perp \neq \{0\}$ and

$$V^\perp \cap \ker(T) \;\subset\; V^\perp \cap H_0 \;\subset\; V^\perp \cap V \;=\; \{0\}\,.$$

Moreover, it is not hard to show that T maps $V^\perp$ into $V^\perp$. Thus the restriction of T to $V^\perp$ would be a nonzero, compact self-adjoint operator of $V^\perp$ and would have at least one nonzero eigenvalue. This would imply by Corrolary 1.3.7 that T has an eigenvector in $V^\perp$ with a nonzero eigenvalue contradicting the fact that

$$V^\perp \cap \cup_{k\geq 1} H_k \;\subset\; V^\perp \cap V \;=\; \{0\}\,.$$

3. For each $k \geq 1$, the finite dimensional subspace H_k possesses a finite orthonormal basis

$$\mathcal{B}_k \;=\; \Big\{ e_{k,1}\,,\; e_{k,2}\,,\ldots\,,\; e_{k,n_k} \Big\}\,.$$

Besides the closed subspace $H_0 = \ker(T)$ of the separable space H, is either null, in which case we set $\mathcal{B}_0 = \emptyset$, or admits an at most countable orthonormal basis $\mathcal{B}_0$.
It follows that

$$\mathcal{B} \;=\; \cup_{k=0}^{+\infty} \mathcal{B}_k$$

is a countable orthonormal basis of eigenvectors of T. Furthermore

$$Tx \;=\; \sum_{k=1}^{+\infty} \sum_{j=1}^{n_k} \mu_k \langle x,\, e_{k,j} \rangle e_{k,j} \qquad \text{for all} \;\; x \in H\,.$$

Constructive proof (Sketchy)
We shall prove again this Theorem 1.4.1 by successive applications of Corollary 1.3.7.

Let $H_1 = H$ assumed to be nontrivial and set $T_1 = T$.
By the second part of Corollary 1.3.7,
there exist an eigenvalue λ_1 of T_1 and a corresponding eigenvector φ_1 such that $||\varphi_1|| = 1$ and $|\lambda_1| = ||T_1||$. Set $H_2 := \{\varphi_1\}^\perp$. Thus H_2 is a closed subspace of H_1 and $T(H_2) \subset H_2$ (i.e H_2 is T-invariant).
Now let T_2 be the restriction of T to H_2. Then T_2 is compact self adjoint operator in $\mathcal{B}(H_2)$
If $T_2 \neq 0$,then there exists an eigenvalue λ_2 of T_2 and corresponding eigenvector φ_2 such that $||\varphi_2|| = 1$ and $|\lambda_2| = ||T_2|| \leq ||T_1|| = |\lambda_1|$
$\{\varphi_1, \varphi_2\}$ is orthonormal.

$$H_3 = \{\varphi_1, \varphi_2\}^\perp$$

H_3 is a closed subspace of H and $TH_3 \subset H_3$

Letting T_3 be the restriction of T to H_3, we have that T_3 is a compact self-adjoint operator in $\mathcal{B}(H_3)$. Continuing in this manner, the process stops when $T_n = 0$ or else we get a sequence $\{\lambda_n\}$ of eigenvalues of T and corresponding orthonormal set $\{\varphi_1, \varphi_2, \varphi_3...\}$ of eigenvectors such that

$$|\lambda_{n+1}| = ||T_{n+1}|| \leq ||T_n|| = |\lambda_n| \qquad n = 1, 2, 3... \tag{1.4.1}$$

Claim : If $\{\lambda_n\}$ is an infinite sequence, then $\lambda_n \longrightarrow 0, \quad n \longrightarrow \infty$.
Proof of Claim. Suppose by contradiction , there exist $\epsilon > 0$ such that $|\lambda_n| \geq \epsilon$ for all $n \in \mathbb{N}$
Hence for $n \neq m$, we have that,

$$||T\varphi_n - T\varphi_m||^2 = ||\lambda\varphi_n - \lambda\varphi_m||^2 = \lambda_n^2 + \lambda_m^2 > \epsilon \tag{1.4.2}$$

But this is impossible, since $\{T\varphi_n\}$ has a convergent subsequence due to the compactness of T. We therefore conclude that $\lambda_n \longrightarrow 0, n \longrightarrow \infty$. $\square$

Now, we prove the representation of T as asserted in the theorem.
Case I. $T_n = 0$ for some n

$$x_n := x - \sum_{k=1}^{n} < x, \varphi_k > \varphi_k$$

It is evident that x_n is orthogonal to φ_i for $1 \leq i \leq n$
Therefore, $x_n \in H_n$

$$0 = T_n x_n = Tx - T(\sum_{k=1}^{n} \langle x, \varphi_k \rangle \varphi_k)$$

$$\Rightarrow Tx = \sum_{k=1}^{n} \lambda_k \langle x, \varphi_k \rangle \varphi_k$$

Case II. $T_n \neq 0$ for all $n \in \mathbb{N}$

$$\begin{aligned} ||Tx - \sum_{k=1}^{n} \lambda_k \langle x, \varphi_k \rangle \varphi_k|| \; = \; ||T_n x_n|| \leq \quad & ||T_n|| \, ||x_n|| \\ = \quad & |\lambda_n| \, ||x_n|| \\ \leq \quad & |\lambda_n| ||x|| \longrightarrow 0 \end{aligned}$$

$$\Rightarrow ||Tx - \sum_{k=1}^{n} \lambda_k \langle x, \varphi_k \rangle \varphi_k|| \quad \text{as} \longrightarrow 0, \quad n \longrightarrow \infty.$$

Hence, $\quad Tx \; = \; \sum_{k=1}^{\infty} \lambda_k \langle x, \varphi_k \rangle \varphi_k \, .$ $\square$

Chapter 2

Bilinear Maps and Forms

2.1 Bilinear maps

Definition 2.1.1 (Bilinear maps)

Let E, F and G be three arbitrary vector spaces over $\mathbb{K}$.
A **bilinear map** Φ from $E \times F$ into G is a mapping $\Phi : E \times F \longrightarrow G$ satisfying the following two conditions :

(i) $\Phi(\alpha_1 x_1 + \alpha_2 x_2, y) = \alpha_1 \Phi(x_1, y) + \alpha_2 \Phi(x_2, y)$ for all $x_1, x_2 \in E$, $y \in F$ and $\alpha_1, \alpha_2 \in \mathbb{K}$.

(ii) $\Phi(x, \alpha_1 y_1 + \alpha_2 y_2) = \alpha_1 \Phi(x, y_1) + \alpha_2 \Phi(x, y_2)$ for all $x \in E$, $y_1, y_2 \in F$ and $\alpha_1, \alpha_2 \in \mathbb{K}$.

This means that Φ is separately linear with respect to each of its two arguments (variables).

When $E = F$, a bilinear map from $E^2 = E \times E$ into G is called a G-valued **bilinear map on** E.

Remark 2.1.2

Note that the above two conditions that define the bilinearity of Φ are also respectively equivalent to the following :

(i) $\Phi(x_1 + \alpha x_2, y) = \Phi(x_1, y) + \alpha \Phi(x_2, y)$ for all $x_1, x_2 \in E$, $y \in F$ and $\alpha \in \mathbb{K}$.

(ii) $\Phi(x, y_1 + \alpha y_2) = \Phi(x, y_1) + \alpha \Phi(x, y_2)$ for all $x \in E$, $y_1, y_2 \in F$ and $\alpha \in \mathbb{K}$.

There are many interesting bilinear maps in the literature. Let's us mention few ones.

Examples 2.1.3

1. Given a $\mathbb{K}$-vector space V, the scalar multiplication defined from $\mathbb{K} \times V$ into V as
$$(\alpha,\, v) \;\longmapsto\; \alpha \cdot v$$
is bilinear.

2. Given a $\mathbb{K}$-vector space V and denoting its space of endomorphisms by $\mathscr{L}(V)$, then the composition mapping defined from $\mathscr{L}(V) \times \mathscr{L}(V)$ into $\mathscr{L}(V)$ as
$$(f,\, g) \;\longmapsto\; f \circ g$$
is bilinear.

3. Given a $\mathbb{K}$-vector space V and denoting its space of endomorphisms by $\mathscr{L}(V)$, then the "valuation" mapping defined from $\mathscr{L}(V) \times V$ into V as
$$(\varphi,\, v) \;\longmapsto\; \varphi(v)$$
is bilinear.

   In particular, by considering a $\mathbb{K}$-vector space E with algebraic dual E^*, then the duality pairing defined from $E^* \times E$ into $\mathbb{K}$ as
$$(f,\, x) \;\longmapsto\; \langle f,\, x\rangle := f(x)$$
is bilinear.

4. Let V_1 and V_2 be two arbitrary vector spaces over the same field and let $f_i \in \mathscr{L}(V_i)$, $i = 1,\, 2$ be fixed. Then the mapping defined from $V_1 \times V_2$ into $V_1 \times V_2$ as
$$(x, y) \;\longmapsto\; \big(f_1(x),\, f_2(y)\big)$$
is bilinear.

5. Let V_1 and V_2 be two arbitrary vector spaces over the same field $\mathbb{K}$. Let $f \in V_1^*$, $g \in V_2^*$ be fixed. Then the mapping defined from $V_1 \times V_2$ into $\mathbb{K}$ as
$$(x, y) \;\longmapsto\; f_1(x) f_2(y)$$
is bilinear.

6. Given a real prehilbertian space $\big(V,\, \langle\, ,\, \rangle\big)$, the inner product $\langle\, ,\, \rangle$ is bilinear.

   In particular, the following mappings are bilinear:

a) For any fixed natural number n, the *dot product* on $\mathbb{R}^n$, that is the mapping defined from $\mathbb{R}^n \times \mathbb{R}^n$ into $\mathbb{R}$ as

$$(x,\, y) \;\longmapsto\; x \cdot y \;=\; \sum_{k=1}^{n} x_k y_k$$

$$\text{with} \quad x = (x_1,\, \ldots,\, x_n) \text{ and } y = (y_1,\, \ldots,\, y_n)\,,$$

is bilinear.

But the mapping defined from $\mathbb{C}^n \times \mathbb{C}^n$ into $\mathbb{C}$ as

$$(x,\, y) \;\longmapsto\; \sum_{k=1}^{n} x_k \overline{y}_k \qquad \text{with} \quad x = (x_1,\, \ldots,\, x_n) \text{ and } y = (y_1,\, \ldots,\, y_n)\,,$$

is not bilinear.

b) The trace-form on the real vector space $\mathbb{M}_n(\mathbb{R})$ of square matrices of order n, defined as

$$(A, B) \;\longmapsto\; \operatorname{tr}\left(AB^T\right)\,,$$

is a bilinear form.

c) The mapping defined from $\mathcal{C}^1([0,1]) \times \mathcal{C}^1([0,1])$ into $\mathbb{R}$ as

$$(f,\, g) \;\longmapsto\; \int_0^1 f'(x) g'(x)\, dx$$

is bilinear.

7. Given a natural number $n \geq 2$ and two natural numbers p and q such that $p + q = n$, the mapping $\langle \cdot\,,\, \cdot \rangle_{p,q}$ defined from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$ by

$$\begin{aligned} \langle x,\, y \rangle_{p,q} &= \textstyle\sum_{k=1}^{p} x_k y_k \;-\; \sum_{k=1}^{q} x_{p+k} y_{p+k} \\ &= x_1 y_1 + \ldots + x_p y_p - x_{p+1} y_{p+1} - \ldots - x_n y_n\,, \end{aligned}$$

$$\text{with} \quad x = (x_1,\, \ldots,\, x_n) \text{ and } y = (y_1,\, \ldots,\, y_n)\,,$$

is bilinear.

8. Let $\mathfrak{g}$ be a finite dimensional vector space over $\mathbb{R}$ equipped with a bilinear multiplication $[\,\cdot\,,\, \cdot\,]$ from $\mathfrak{g} \times \mathfrak{g}$ to $\mathfrak{g}$ satisfying furthermore

$$[x,\, x] \;=\; 0\,, \qquad \forall x \in \mathfrak{g}$$

and

$$[[x, y],\, z] \;+\; [[y, z],\, x] \;+\; [[z, x],\, y] \;=\; 0 \quad \forall x,\, y,\, z \in \mathfrak{g} \qquad \text{(Jacobi's Identity).}$$

Then $\mathfrak{g}$ is called *Lie Algebra*. For more details see [14]

Proposition 2.1.4

Given three arbitrary vector spaces E, F and G over $\mathbb{K}$, every bilinear map Φ from $E \times F$ into G can be identified with a linear map $\Phi^{\#}$ from E into $\mathscr{L}(F, G)$; the space of linear maps from F into G, which takes each $x \in E$ to the linear map $y \mapsto \Phi(x, y)$.
In other words if we denote by $\mathscr{L}(E,\, F,\, G)$ the set of bilinear maps from $E \times F$ into G, we have

$$\Phi \in \mathscr{L}(E,\, F;\, G) \quad \Longleftrightarrow \quad \Phi^{\#} \in \mathscr{L}\big(E,\, \mathscr{L}(F,\, G)\big)$$

where for each $x \in E$, $\Phi^{\#}(x)$ is defined from F into G by

$$\big[\Phi^{\#}(x)\big]\,(y) \;:=\; \Phi(x, y) \quad \forall y \in F\,.$$

In the next we introduce bilinear forms on which we shall focus our attention in the sequel.

2.2 Bilinear Forms and Spaces

Definition 2.2.1 (Bilinear form)

A **bilinear form** f acting on a vector space V over the field $\mathbb{K}$ is a $\mathbb{K}$-valued bilinear map on V. This means that f is defined from $V^2 := V \times V$ into $\mathbb{K}$;

$$(u, v) \;\longmapsto\; f(u, v)\,,$$

and is linear in its first and second arguments. That is; for all $\lambda \in \mathbb{K}$,

(i) $f(u_1 + \lambda u_2,\, v) \;=\; f(u_1,\, v) \,+\, \lambda f(u_2,\, v)$, for all $u_1,\, u_2,\, v \in V$, and moreover

(ii) $f(u,\, v_1 + \lambda v_2) \;=\; f(u,\, v_1) \,+\, \lambda f(u,\, v_2)$, for all $u,\, v_1,\, v_2 \in V$.

A vector space V equipped with a bilinear form f is called a *bilinear space* and is denoted by $(V,\, f)$.

Proposition 2.2.2

1. Given a natural number n, the real vector space $\mathbb{R}^n$ equipped with the dot product is a bilinear space called *euclidean space*.

2. Given a natural number $n \geq 2$ and two natural numbers p and q such that $p+q=n$, the real vector space $\mathbb{R}^n$ equipped with the bilinear form $\langle \cdot, \cdot \rangle_{p,q}$ defined by

$$\langle x, y \rangle_{p,q} \;=\; x_1 y_1 + \ldots + x_p y_p - x_{p+1} y_{p+1} - \ldots - x_n y_n \,,$$
$$\text{for} \quad x = (x_1, \ldots, x_n) \text{ and } y = (y_1, \ldots, y_n) \;\text{ in }\; \mathbb{R}^n \,,$$

is a bilinear space called *pseudo-euclidean space* et denoted by $\mathbb{R}^{p,q}$.

3. Let $\mathfrak{g}$ be a finite dimensional Lie Algebra over $\mathbb{R}$. Define with the help of the Lie bracket $[\cdot, \cdot]$, for each $x \in \mathfrak{g}$, the adjoint endomorphism $\mathrm{ad}(x) = [x, \cdot]$ on $\mathfrak{g}$, that is, $\mathrm{ad}(x)(y) = [x, y]$ for all $y \in \mathfrak{g}$.
Then the bilinear map defined by

$$B(x, y) \;=\; \mathrm{tr}\,[\mathrm{ad}(x), \mathrm{ad}(y)] \qquad \forall\, x,\, y \in \mathfrak{g}\,,$$

is a bilinear form called the *Killing form* of $\mathfrak{g}$.

If $\mathfrak{g} = \mathfrak{gl}_n(\mathbb{R})$, then its Killing form is just the trace form (See Example ...).

For more details, the interested reader is refered to the book by Serre [14].

Definitions 2.2.3 (Classification of Bilinear forms)

Let f be a bilinear form on a $\mathbb{K}$-vector space V. Then:

- f is **symmetric** if

$$f(x, y) \;=\; f(y, x)\,, \qquad \forall\, x,\, y \in V.$$

For instance the dot product on $\mathbb{R}^n$ is a symmetric bilinear form. The trace-form on the vector space of matrices $\mathbb{M}_n(K)$ is also symmetric.

- f is **skew-symmetric** or **anti-symmetric** if

$$f(x, y) \;=\; -f(y, x)\,, \qquad \forall\, x,\, y \in V.$$

For instance the mapping defined from $\mathbb{K}^2 \times \mathbb{K}^2$ into $\mathbb{K}$ as

$$\Big((x_1, x_2)\,;\; (y_1, y_2)\Big) \;\longmapsto\; x_1 y_2 - x_2 y_1 \;=\; \det\big((x_1, x_2)\,;\, (y_1, y_2)\big)$$

is a skew-symmetric bilinear form.

- f is **alternating** if

$$f(x, x) \;=\; 0\,, \qquad \forall\, x \in V.$$

Note that every alternating form is skew-symmetric (To verify this just expand $f(x+y,\, x+y)$). The converse is true for any field of characteristic different from 2, for instance $\mathbb{K}$ of which characteristic is 0.

- f is **nondegenerate** if

  $$\text{for any } x \in V, \quad \text{whenever} \quad f(x,y) = 0 \ \text{ for all } y \in V, \quad \text{then} \quad x = 0\,.$$

  That is $f^{\#}$ is one-to-one; meaning also that the $\ker(f^{\#})$ (also called the *radical of the bilinear form* f) is null.

  In the finite dimensional case, f is nondegenerate if and only if $f^{\#}$ is an isomorphism.

  Beware in any case that the existence of a nonzero vector v satisfying $f(v,v) = 0$, does not imply that f is degenerate!

- f is said to be **positive definite** if f is nondegenerate and $f(x,x) \geq 0$ for all $x \in V$. Equivalently, f is positive definite if and only if $f(x,x) > 0$ for all nonzero $x \in V \setminus \{0\}$.

  f is said to be **positive semidefinite** if $f(x,x) \geq 0$ for all $x \in V$.

  f is said to be **negative definite** if f is nondegenerate and $f(x,x) \leq 0$ for all $x \in V$. Equivalently, f is negative definite if and only if $f(x,x) < 0$ for all nonzero $x \in V \setminus \{0\}$.

  f is said to be **negative semidefinite** if $f(x,x) \leq 0$ for all $x \in V$.

  Otherwise, f is said to be **indefinite**.
  In this case the function $x \mapsto f(x,x)$ changes sign.

- When V is a normed space, f is said to be **coercive (or elliptic)** if there exists a positive constant β such that

  $$f(x,x) \geq \beta ||x||_V^2 \qquad \text{for all } x \in V.$$

- When V is a normed space, f is said to be **bounded** if there exists a positive constant real number M such that

  $$|f(x,y)| \ \leq \ M||x||\,||y|| \qquad \forall\, x,\, y \in V\,.$$

Proposition 2.2.4

Let f be a symmetric bilinear form on a normed space V over the field $\mathbb{K}$. Then f is identically zero if and only if f vanishes on the diagonal of $V \times V$; that is,

$$\Big(f(x,x) \ = \ 0 \quad \forall x \in V\Big) \quad \Longleftrightarrow \quad f = 0 \text{ on } V \times V\,.$$

Proof. If f is zero on $V \times V$, it is trivial that f also vanishes on the diagonal of $V \times V$.
Now suppose that f vanishes on the diagonal of $V \times V$; that is $f(a, a) = 0$ for all $a \in V$. Therefore, for any $x,\, y \in V$, we have

$$f(x, y) \;=\; \frac{1}{2}\,[f(x + y,\, x + y) \;-\; f(x,\, x) \;-\; f(y,\, y)] \;=\; 0\,.$$

We are done. □

Proposition 2.2.5 (Continuity of bounded bilinear forms)

Let f be a bilinear form on a normed space V over the field $\mathbb{K}$.

1. f is continuous if and only if f is bounded.

2. When V is a Banach space, f is continuous if and only if f is separately continuous with respect to its two variables (i.e., the functions $x \mapsto f(x, v)$ and $y \mapsto f(u, y)$ are continuous for all $u,\, v \in V$).

   This result does not hold when V is not a Banach space. *(To see this it suffices to consider the subspace c_{oo} of l^∞)* consisting of sequences having finitely many nonzero terms and the bilinear form

$$\varphi(u,\, v) \;=\; \sum_{k=1}^{+\infty} u_k v_k\,, \qquad u,\, v \in c_{oo}\,.$$

Proof.

1. If f is bounded, then it is immediate that f is continuous because for any $(a, b) \in V^2$ fixed, we have

$$\begin{aligned} |f(x, y) \;-\; f(a, b)| \;&=\; |f(x,\, y - b)| \;+\; f(x - a,\, b)| \\ &\leq\; M||x||\,||y - b|| \;+\; M||x - a||\,||b|| \end{aligned}$$

   showing that f is locally Lipschitz.

   Conversely, if f is continuous, then it is continuous at $(0, 0)$. And so there exists a real number $\delta > 0$ such that

$$\forall (u, v) \in V^2, \quad ||u|| < \delta \text{ and } ||u|| < \delta \quad \Longrightarrow \quad |f(u, v)| < 1\,.$$

   Therefore

$$|f(x, y)| \;\leq\; \frac{4}{\delta^2}||x||\,||y|| \quad \forall x,\, y \in V\,.$$

2. Suppose that V is a Banach space.
If f is continuous, then it is clearly separately continuous.
Now suppose that it is separately continuous. Then $\{f(x,\,\cdot)\}_{||x||\leq 1}$ is a family of bounded linear forms of the Banach space V such that for each $y \in V$, the set $\{f(x,y)\}_{||x||\leq 1}$ is bounded since the map $x \longmapsto f(x,y)$ is continuous. Therefore, by the Uniform Boundedness Principle (Theorem 1.1.22) we have

$$M \;:=\; \sup_{||y||\leq 1}\left(\sup_{||x||\leq 1} ||f(x,y)||\right) \;<\; +\infty\,.$$

It follows that

$$||f(x,y)|| \leq\; M||x||\,||y||\,, \quad \forall\, x,\, y \in V\,,$$

and so the bilinear form f is bounded.

Proposition 2.2.6 (Smoothness of bounded bilinear forms)

Every bounded bilinear form f on a normed space V over the field $\mathbb{K}$ is of class $\mathcal{C}^\infty$.

The Fréchet derivative of f at a given point $(u_0, v_0) \in V \times V$ is $f'(u_0, v_0) : V \times V \to V$ defined by

$$\Big[f'(u_0,v_0)\Big](h,k) \;=\; f(h,v_0) \,+\, f(u_0,k)\;, \qquad \forall (h,k) \in V \times V\,.$$

Proof.
Let $(u_0, v_0) \in V \times V$. Then the mapping

$$(h,k) \;\longmapsto\; f(h,v_0) \,+\, f(u_0,k)$$

is linear and bounded from $V \times V$ to $\mathbb{K}$ according to the boundedness of the bilinear form f. Moreover we have for all $(h,k) \in V \times V$,

$$f(u_0+h, v_0+k) \;-\; f(u_0,v_0) \;-\; \big(f(h,v_0) \,+\, f'(u_0,k)\big) \;=\; f(h,k)$$

with

$$|f(h,k)| \;\leq\; M||h||\,||k|| \;=\; o\Big(||(h,k)||^2\Big)\,.$$

Proposition 2.2.7 (Decomposition of a bilinear form)

Every bilinear form $f : V \times V \to \mathbb{R}$ has a unique decomposition $f \;=\; f_1 + f_2$ as the sum of a symmetric bilinear form f_1 and a skew-symmetric bilinear form f_2. These forms f_1 and f_2 are respectively defined by

$$\begin{aligned} f_1(x,y) \;&=\; \tfrac{1}{2}\big[f(x,y) \,+\, f(y,x)\big] \\ \text{and} \\ f_2(x,y) \;&=\; \tfrac{1}{2}\big[f(x,y) \,-\, f(y,x)\big]\,. \end{aligned}$$

Proof.
Let f be a bilinear form on V.

- *Uniqueness of the decomposition.* Suppose that $f = g + h$; where g is a symmetric bilinear form and h a skew-symmetric bilinear form. Then for all x, $y \in V$, we have:

$$\begin{aligned} f(x,y) &= g(x,y) + h(x,y) \\ \text{and} \quad f(y,x) &= g(y,x) - h(x,y) \quad \text{since } h(y,x) = -h(x,y)\,. \end{aligned}$$

  Therefore, on the one hand, $f(x,y) + f(y,x) = 2g(x,y)$ that implies

$$g(x,y) = \frac{1}{2}\big[f(x,y) + f(y,x)\big]\,,$$

  and on the other hand, $f(x,y) - f(y,x) = 2h(x,y)$ that implies

$$h(x,y) = \frac{1}{2}\big[f(x,y) - f(y,x)\big]\,.$$

- *Existence.* By setting

$$\begin{aligned} f_1(x,y) &= \tfrac{1}{2}\big[f(x,y) + f(y,x)\big] \\ \text{and} \quad & \\ f_2(x,y) &= \tfrac{1}{2}\big[f(x,y) - f(y,x)\big]\,, \end{aligned}$$

  it is evident that f_1 and f_2 are bilinear forms which are respectively symmetric and skew-symmetric and satisfy furthermore $f_1 + f_2 = f$.

It follows that given an arbitrary bilinear form f, there corresponds a symmetric part $f^\star$ defined by:

$$f^\star(x,y) = \frac{f(x,y) + f(y,x)}{2}\,. \tag{2.2.1}$$

Observe that f is symmetric if and only if $f^\star = f$.

2.3 Notions of Orthogonality. Orthogonal bases

Definitions 2.3.1

Let V be a real vector space and $b : V \times V \to \mathbb{R}$ a symmetric bilinear form.

- We say that two vectors $u,\, v \in V$ are *b-orthogonal* (or simply *orthogonal*) if $b(u, v) = 0$. And in this case we write $u \perp_b v$ or simply $u \perp v$ when there is no ambiguity.

  Given a nonempty subset S of V, the set

  $$S^{\perp} \;=\; \big\{ u \in V \;:\; b(u, v) = 0 \;\; \forall v \in S \,\big\}$$

  is called the *b-orthogonal complement* of S; it is the collection of all vectors of V that are b-orthogonal to every vector of S.
  Observe that $S^{\perp}$ is always a linear subspace of V, whether the nonempty set S is a subspace or not.
  For $S = \{a\}$, we abbreviate $\{a\}^{\perp}$ by $a^{\perp}$.

- The b-orthogonal complement of the whole space V; $V^{\perp}$, is also called the *radical* of b and is denoted by $\mathrm{rad}(b)$.

- A vector $v \in V$ is called *b-isotropic*, if $b(v, v) = 0$.
  The set of all isotropic vectors of V is called the *b-isotropic cone* of V. An isotropic cone might not be convex (and so it might not be a vector subspace).
  Always beware that the existence of a b-isotropic vector does not imply that b is degenerate. For instance $(1,\, 1)$ is isotropic for the non degenerate bilinear form $b : \mathbb{R}^2 \times \mathbb{R}^2 \longrightarrow \mathbb{R}$ defined by

  $$b\Big((x_1,\, x_2);\, (y_1,\, y_2) \Big) \;=\; x_1 y_1 \,-\, x_2 y_2 \,.$$

  Note that a vector v which is not b-isotropic (i.e., $b(v, v) \neq 0$) is said to be *b-anisotropic*.

Definition 2.3.2 (Orthogonal basis)

Given a n-dimensional nontrivial vector space V and a symmetric bilinear form b on V, we say that a basis $\mathscr{B} = (u_1\,, \ldots,\, u_n)$ of V is *b-orthogonal* if any two distinct basis vectors are b-orthogonal; i.e.,

$$f(u_i, u_j) \;=\; 0 \quad \text{for all} \;\; i,\, j = 1, \ldots,\, n \;\; \text{with } i \neq j \,.$$

Proposition 2.3.3

Let $b : V \times V \to \mathbb{R}$ be a symmetric bilinear form and u a non-isotropic vector of the bilinear space $(V,\, b)$. Then we have the direct sum

$$V \;=\; \langle u \rangle \oplus \{u\}^{\perp} \,.$$

Proof.
On the one hand, we have

$$\langle u\rangle \cap \left(u^{\perp}\right) \;=\; \mathrm{Span}(u) \cap \{u\}^{\perp} \;=\; \{0\}$$

because u is non-isotropic.
On the other hand, we have for any $x \in V$, $x = x_1 + x_2$ with

$$x_1 \;=\; \frac{b(u,x)}{b(u,u)}u \;\in\; \langle u\rangle \qquad \text{and} \quad x_2 \;=\; x - \frac{b(u,x)}{b(u,u)}u \;\in\; \{u\}^{\perp}\,.$$

□

Definition 2.3.4 (Fourier coeficient)

Let $b:\, V \times V \;\to\; \mathbb{R}$ be a symmetric bilinear form and u a non-isotropic vector. For any vector $x \in V$, the scalar $\frac{b(u,x)}{b(u,u)}$ is called the *Fourier coefficient* of x *with respect to* u and is denoted by $c_u(x)$.
The vector $x_1 = c_u(x)u$ is called the *b-orthogonal projection* of x *in the direction* of u.
Note that c_u defines a linear map (more precisely a linear form) from V onto $\mathbb{R}$.

Theorem 2.3.5 (Existence of Orthogonal Bases) [4]

Let b be a symmetric bilinear form on a n-dimensional nontrivial vector space V over $\mathbb{K}$.
Then V has a b-orthogonal basis.

Proof. By induction on the dimension $n \geq 1$.
For $n = 1$, the result is trivially true.
Suppose that for some natural number n, every symmetric bilinear space $(W,\, B)$ with $\dim W = n$ has a B-orthogonal basis.
Let $(V,\, b)$ be a symmetric bilinear space such that $\dim V = n+1$. Let us show that $(V,\, b)$ has a b-orthogonal basis. If all the vectors of V are b-isotropic, then by Proposition 2.2.4 $b = 0$ on V, and in this case any basis of V is well b-orthogonal (i.e., 0-orthogonal). Otherwise, there exists at least one anisotropic vector u_1 in $(V,\, b)$. But by Proposition 2.3.3, we have

$$V \;=\; \mathbb{K}u_1 \oplus \{u_1\}^{\perp}\,.$$

Setting

$$W \;=\; \{u_1\}^{\perp} \qquad \text{and} \qquad B(x,y) \;=\; b(x,y)\;\forall x,\, y \in W\,,$$

it is clear that $(W,\, B)$ is a symmetric bilinear space such that $\dim W = n$. Therefore by the induction hypothesis, $(W,\, B)$ has a B-orthogonal basis, say

$(u_2, \ldots, u_{n+1})$. Hence it is clear that $(u_1, u_2, \ldots, u_{n+1})$ is a b-orthogonal basis of V since

$$V = \mathbb{K}u_1 \oplus W \qquad \text{with} \quad W = \{u_1\}^{\perp_b} \text{ and } B = b\Big|_{W \times W}.$$

The proof is completed. □

2.4 Isometries and Similarities of a nondegenerate bilinear form

Let Φ be a nondegenerate bilinear form on a real vector space V.

Definition 2.4.1 (Isometries)

An *isometry* of the bilinear space (V, Φ) is an endomorphism g of V which preserves Φ; that is, a linear map $g : V \to V$ such that

$$\Phi\big(g(x), g(y)\big) = \Phi(x, y), \qquad \forall\, x, y \in V.$$

We usually think of an isometry as preserving inner products and norms, and so preserving distances and angles.

Definition 2.4.2 (Similarities)

A *similarity* of of the bilinear space (V, Φ) is an endomorphism g of V which preserves such that

$$\Phi\big(g(x), g(y)\big) = \lambda_g \Phi(x, y), \qquad \forall\, x, y \in V;$$

where λ_g is a scalar depending on g but not on the variables x and y.
Similarities preserve angles and allow change of scales.

2.5 Matrix representation and Diagonization Theorem

In this section we shall give the matrix representation of a bilinear form on a finite dimensional space with a prescribed basis. The diagonization theorem will also be given.
Let V be an n-dimensional space over $\mathbb{K}$ and $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ be a basis of V.

Theorem 2.5.1

A function $f : V \times V \to \mathbb{K}$ is bilinear if and only if there is a square matrix $\mathbf{A}$ of order n and of elements of $\mathbb{K}$ such that

$$f(x, y) \;=\; [x]_{\mathscr{B}}^T \, \mathbf{A} \, [y]_{\mathscr{B}} \,, \quad \forall x,\, y \in V \,;$$

where $[x]_{\mathscr{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $[y]_{\mathscr{B}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$ are respectively the column-matrices made of the components of x and y in the basis $\mathscr{B}$; that is

$$x \;=\; \sum_{i=1}^{n} x_i e_i \quad \text{and} \quad y = \sum_{j=1}^{n} y_i e_j \,.$$

In this case, the matrix $\mathbf{A}$ is unique and called the *matrix representation of the bilinear form f* with respect to $\mathscr{B}$.

This matrix is denoted by $[f]_{\mathscr{B}}$ and we have

$$[f]_{\mathscr{B}} = \begin{pmatrix} f(e_1, e_1) & \dots & f(e_1, e_n) \\ \vdots & \ddots & \vdots \\ f(e_n, e_1) & \dots & f(e_n, e_n) \end{pmatrix} .$$

Proof.

• Let f be a bilinear form on V and set

$$[f]_{\mathscr{B}} \;=\; \begin{pmatrix} f(e_1, e_1) & \dots & f(e_1, e_n) \\ \vdots & \ddots & \vdots \\ f(e_n, e_1) & \dots & f(e_n, e_n) \end{pmatrix} .$$

For any x, $y \in V$, express x and y as linear combinations of the elements of the basis $\mathscr{B}$. That is:

$$x \;=\; \sum_{i=1}^{n} x_i e_i \quad \text{and} \quad y \;=\; \sum_{j=1}^{n} y_i e_j \,.$$

Then we have

$$\begin{aligned} f(x,y) &= f\left(\sum_{i=1}^{n} x_i e_i\,,\ \sum_{j=1}^{n} y_j e_j\right) \\ &= \sum_{i=1}^{n} x_i\, f\left(e_i\,,\ \sum_{j=1}^{n} y_j e_j\right) \\ &= \sum_{i=1}^{n} x_i \left(\sum_{j=1}^{n} y_j f(e_i, e_j)\right) \\ &= [x]_{\mathscr{B}}^T\, [f]_{\mathscr{B}}\, [y]_{\mathscr{B}} \end{aligned}$$

since $[x]_{\mathscr{B}} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$ and $[y]_{\mathscr{B}} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}$

• Conversely, given a square matrix $A = (a_{ij})_{\substack{1 \le i \le n \\ 1 \le j \le n}}$ of order n, the mapping φ defined from $V \times V$ into $\mathbb{K}$ by

$$\varphi(x,y) \;=\; [x]_{\mathscr{B}}^T\, A\, [y]_{\mathscr{B}}$$

is bilinear and we have $\varphi(e_i, e_j) = a_{ij}$ for all i, j belonging to $\{1, 2, \ldots, n\}$. □

Corollary 2.5.2 (Expression of a bilinear form on $\mathbb{K}^n$)

The general analytic expression of a bilinear form f on $\mathbb{K}^n$ is

$$f(x,y) \;=\; \sum_{i,j=1}^{n} a_{i\,j} x_i y_j$$

$$\text{for all} \quad x = (x_1, \ldots, x_n)\,;\ y = (y_1, \ldots, y_n) \in \mathbb{K}^n\,;$$

and where a_{ij} are constant scalars.

Note that by denoting the canonical basis of $\mathbb{K}^n$ by $\big(e_i\big)_{1 \le i \le n}$, we have in fact $a_{ij} = f(e_i, e_j)$ for all i, j.

Corollary 2.5.3 (Smoothness)

Every bilinear form on a finite dimensional normed space is of class $\mathcal{C}^\infty$.

Proof. Follows from Corollary 2.5.2 and the fact that every polynomial (of several indeterminates) is of class $\mathcal{C}^\infty$. □

Theorem 2.5.4 (Change of basis formula in the finite dimensional case)

Let f be a bilinear form on a finite dimensional vector space V over $\mathbb{K}$. Let $\mathscr{B}$ and $\mathscr{C}$ be two bases for V and let P be the transition matrix from $\mathscr{B}$ to $\mathscr{C}$. Then

$$[f]_{\mathscr{C}} \;=\; P^T[f]_{\mathscr{B}}P\,.$$

Proof. Let $u,\, v \in V$ with

$$\begin{aligned} [u]_{\mathscr{B}} &= U, \qquad [v]_{\mathscr{B}} = V, \\ [u]_{\mathscr{C}} &= X, \qquad [v]_{\mathscr{C}} = Y\,. \end{aligned}$$

Let $A = (a_{i,j})$ be the matrix representing f with respect to $\mathscr{B}$. i.e

$$[f]_{\mathscr{C}} \;=\; A\,.$$

Now,

$$U \;=\; PX \quad \text{and} \quad V \;=\; PY$$

since P is the transition matrix from $\mathscr{B}$ to $\mathscr{C} = (w_1,\, w_2,\, \ldots,\, w_n)$

$$\begin{aligned} f(u,v) &= [u]_{\mathscr{B}}^T[f]_{\mathscr{B}}[v]_{\mathscr{B}} \\ &= U^TAV \\ &= (PX)^TA(PY) \\ &= (X^TP^T)A(PY) \\ &= X^T(P^TAP)Y \\ &= [u]_{\mathscr{C}}(P^TAP)[v]_{\mathscr{C}} \end{aligned}$$

We have $f(w_i, w_j) \;=\; (P^TAP)_{i,j}$. Hence

$$[f]_{\mathscr{C}} \;=\; P^TAP \;=\; P^T[f]_{\mathscr{B}}P.$$

□

Basing on Theorem 2.5.4, we can define, independently from the bases, the rank of a bilinear form on a finite dimensional vector space as follows:

Definition 2.5.5

Let f be a bilinear form on a finite dimensional vector space V over $\mathbb{K}$. The *rank* of f is the rank of its matrix representation with respect to any (given) basis.

Proposition 2.5.6

Let V be an n-dimensional linear space, $\mathscr{B} = (e_1, e_2, \ldots, e_n)$ be a basis of V and let f be a bilinear form on V.
Then f is *symmetric* (resp. *skew-symmetric*), if and only if its matrix representation $[f]_{\mathscr{B}}$ with respect to $\mathscr{B}$ is symmetric (resp. skew-symmetric).

Proof.
• If f is symmetric (resp. skew-symmetric), then for all $i, j \in \{1, \ldots, n\}$, we have

$$f(e_i, e_j) \;=\; f(e_j, e_i) \qquad \Big(\text{resp. } f(e_i, e_j) \;=\; -f(e_j, e_i)\Big)$$

showing that $[f]_{\mathscr{B}}$ is symmetric (resp. skew-symmetric).
• Conversely, if for all $i, j \in \{1, \ldots, n\}$ we have

$$f(e_i, e_j) \;=\; f(e_j, e_i) \qquad \Big(\text{resp. } f(e_i, e_j) \;=\; -f(e_j, e_i)\Big)\,,$$

then for all $x, y \in V$ with

$$x \;=\; \sum_{i=1}^{n} x_i e_i \quad \text{and} \quad y = \sum_{j=1}^{n} y_i e_j\,,$$

we have

$$\begin{aligned} f(x, y) &= [x]_{\mathscr{B}}^T [f]_{\mathscr{B}} [y]_{\mathscr{B}} \in \mathbb{K} \\ &= \left([x]_{\mathscr{B}}^T [f]_{\mathscr{B}} [y]_{\mathscr{B}}\right)^T \\ &= [y]_{\mathscr{B}}^T [f]_{\mathscr{B}}^T [x]_{\mathscr{B}} \\ &= [y]_{\mathscr{B}}^T [f]_{\mathscr{B}} [x]_{\mathscr{B}} \qquad \Big(\text{resp. } -[y]_{\mathscr{B}}^T [f]_{\mathscr{B}} [x]_{\mathscr{B}}\Big) \\ &= f(y, x) \qquad \Big(\text{resp. } -f(y, x)\Big)\,. \end{aligned}$$

□

Theorem 2.5.7

Let V be an n-dimensional linear space, $\mathscr{B}$ a basis of V and let f be a bilinear form on V.
Then f is nondegenerate if and only if its matrix representation $[f]_{\mathscr{B}}$ with respect to $\mathscr{B}$ is nonsingular; that is, $\operatorname{rank}(A) \;=\; \dim(V)$.

Proof.

The proof of the previous theorem follows from the fact that for any square matrix A of order $n \geq 1$, we have

$$\begin{aligned}\left\{ x \in \mathbb{K}^n \,:\, x^T A y = 0_{\mathbb{K}} \;\; \forall y \in \mathbb{K}^n \right\} &= \left\{ x \in \mathbb{K}^n \,:\, x^T A = 0_{\mathbb{K}^n} \right\} \\ &= \left\{ x \in \mathbb{K}^n \,:\, A^T x = 0_{\mathbb{K}^n} \right\} \\ &= \ker\left(A^T\right)\end{aligned}$$

and $\ker\left(A^T\right) = \{0\}$ if and only if A^T is nonsingular, that is, A is nonsingular. $\square$

Definition 2.5.8 (Congruent matrices)

Two matrices $A,\, B \in \mathbb{M}_n(\mathbb{K})$ are *congruent* if there is an invertible matrix $P \in GL_n(\mathbb{K})$ such that

$$B \;=\; P^T A P\,.$$

It is evident that the congruence is an equivalence relation on $\mathbb{M}_n(\mathbb{K})$.

The following property is a consequence of Theorem 2.5.4.

Proposition 2.5.9

Two matrices $A,\, B \in \mathbb{M}_n(\mathbb{K})$ are congruent if and only if they represent the same bilinear form $f : \mathbb{K}^n \times \mathbb{K}^n \longrightarrow \mathbb{K}$ with respect to some different bases.

2.6 Representation of bounded bilinear forms on real Hilbert spaces

Let H be a real Hilbert space. Then it is evident that for every bounded linear operator $A \in \mathcal{B}(H)$, the form

$$(x,\, y) \;\longmapsto\; \langle x,\, Ay \rangle$$

is bilinear and bounded (with $|\langle x,\, Ay\rangle| \leq ||A||\, ||x||\, ||y||$ for all $x,\, y \in H$).

Conversely we have the following result thanks to the Riesz Representation Theorem.

Theorem 2.6.1

Let $f,\colon\; H \times H \longrightarrow \mathbb{R}$ be a bounded bilinear form. Then there exists a bounded linear operator A such that

$$f(x, y) \;=\; \langle x,\, Ay \rangle\,, \qquad \forall x,\, y \in H\,.$$

Moreover, the operator A is symmetric if and only if f is symmetric.

Proof. (Sketch)
By hypothesis, f is bilinear and bounded. Thereofore there exists a non-negatige constant M such that

$$|f(x,y)| \;\leq\; M||x||\,||y||\,, \quad \forall x,\, y \in H\,.$$

Now fix $y \in H$ and consider the map defined from $H \to \mathbb{K}$ by

$$\psi(x) \;=\; f(x,y)\,, \quad \forall x \in H\,.$$

Then ψ is linear by the bilinearity of f and bounded since

$$\psi(x) \;=\; |f(x,y)| \;\leq\; M||x||\,||y||\,, \quad \forall x \in H\,.$$

Therefore by Riesz Representation Theorem there exists a unique element $z \in H$ such that

$$\psi(x) \;=\; \langle x,\, z\rangle\,, \quad \forall x \in H\,.$$

That is

$$f(x,y) \;=\; \langle x,\, z\rangle\,, \quad \forall x \in H\,.$$

Clearly this unique element z depends (a priori) on y, and so by setting

$$z \;=\; Ay$$

we have a well-defined operator A of H such that

$$f(x,y) \;=\; \langle x,\, Ay\rangle\,, \qquad \forall x,\, y \in H\,.$$

Moreover for every $x \in H$, there holds

$$\begin{aligned} ||Ax||^2 \;=\; \langle Ax,\, Ax\rangle \;&=\; f(Ax,x) \\ &\leq\; M||Ax||\,||x||\,, \end{aligned}$$

and so

$$||Ax|| \;\leq\; M||x|| \qquad \text{weither } x \neq 0 \text{ or not,}$$

showning that $||A|| \leq M$. $\square$

Chapter 3

Quadratic forms

3.1 Generalities on Quadratic Forms and Spaces

Definition 3.1.1

A *quadratic form* on a $\mathbb{K}$-vector space V, is a functional Q on V such that there exists a bilinear form f on V satisfying

$$Q(x) \;=\; f(x,x)\,, \qquad \forall\, x \in V\,.$$

First Properties

Proposition 3.1.2 (Polar form of a quadratic form)

For every quadratic form Q on a $\mathbb{K}$-vector space V, there exists a *unique symmetric bilinear form* φ on V such that

$$Q(x) \;=\; \varphi(x,x)\,, \qquad \forall x \in V\,.$$

This unique symmetric bilinear form φ corresponding to Q is called *the polar form* of Q and can be expressed by

$$\varphi(x,y) \;=\; \frac{1}{2}\Big(Q(x+y) - Q(x) - Q(y)\Big)\,, \qquad \forall x,\, y \in V\,.$$

Consequently, there is a one-to-one correspondance between the class of quadratic forms of a vector space V and the class of symmetric bilinear form on V.

Proof.
Q being a quadratic form, there exists a bilinear form f such that

$$Q(x) \;=\; f(x,x) \qquad \text{for all} \;\; x \in V\,.$$

Thus it is not hard to check that $\varphi = f^*$ (the symmetric part of f, cf. Definition ...) is the unique symmetric bilinear form such that $Q(x) \;=\; \varphi(x,x)$ for all $x \in V$. $\square$

Corollary 3.1.3 (Polarization Identity)

For any quadratic form Q with polar form f, we have the identity

$$f(x,y) \;=\; \frac{1}{2}\Big(Q(x+y)-Q(x)-Q(y)\Big), \qquad \forall x,\, y \in V$$

called *polarization identity.*

This section shows in summary that there is a one-to-one correspondence between quadratic forms and symmetric bilinear forms. And this easily reduces the study of quadratic forms to that of symmetric bilinear forms. Besides by Proposition 2.5.6, on finite dimensional vector spaces, symmetric bilinear forms have symmetric matrix representations which will also narrow down the study to that of symmetric matrices. See Proposition 3.2.1 below, the Introduction and Theorem ... .
The one-to-one correspondence between quadratic forms and symmetric bilinear forms is brought to light by the previous Proposition 3.1.2 and clarified by the Polarization Identity (Corollary 3.1.3). By this result, we see that given a quadratic form, we can get a unique corresponding polar form.
Therefore we have the following diagram :

$$\begin{array}{rcl} \text{Quadratic forms} & \longleftrightarrow & \text{Symmetric bilinear forms} \\ q & \longmapsto & f(u,v) = \frac{1}{2}(q(u+v)-q(u)-q(v)) \\ q(u) = f(u,u) & \longleftarrow & f \end{array}$$

Now we provide more criteria that allow to identify quadratic forms without knowing explicitly their corresponding polar forms.

Theorem 3.1.4

A functional Q on a $\mathbb{K}$-vector space V is a quadratic form if the following two conditions are satisfied:

1. Q is 2-homogeneous, and
2. the map $f : V \times V \longrightarrow \mathbb{K}$ defined by

$$f(x,y) \;=\; \frac{1}{2}\Big(Q(x+y)-Q(x)-Q(y)\Big), \qquad \forall x,\, y \in V$$

   is bilinear.

The next theorem provides us with a nontrivial characterization of nonnegative quadratic form.

Theorem 3.1.5 [31]

Let Q be an arbitrary functional on a $\mathbb{K}$-vector space V.

1. If Q is a quadratic form, then :

   a) $Q(tx) \;=\; t^2 Q(x) \quad$ for all $t \in \mathbb{K}$ and $x \in V$.

   b) $Q(x+y) \;+\; Q(x-y) \;=\; 2Q(x) + 2Q(y) \quad$ for all $x,\, y \in V$.

2. Conversely, if Q is nonnegative and satisfies the above two conditions (a) and (b) above, then Q is a nonnegative quadratic form.

Definition 3.1.6

Given a quadratic form Q on a vector space V, the pair $(V,\, Q)$ is called a *quadratic space*.

For instance, every n-dimensional Euclidean space ($\mathbb{R}^n$) is a quadratic space with the following quadratic form

$$Q(x_1,\, \ldots,\, x_n) \;=\; x_1^2 \,+\, \ldots \,+\, x_n^2\,, \qquad \forall\, (x_1,\, \ldots,\, x_n) \in \mathbb{R}^n\,.$$

Definitions 3.1.7 (Classification of Quadratic Forms)

Quadratic forms are classified with respect to the classification of their corresponding polar forms. More precisely, given a quadratic form Q on a $\mathbb{K}$-vector space V,

- Q is **positive definite** if :
  $Q(x) \geq 0$ for all $x \in V$, and $Q(x) = 0$ if and only if $x = 0$.

- Q is **negative definite** if :
  $Q(x) \leq 0$ for all $x \in V$, and $Q(x) = 0$ if and only if $x = 0$.

- Q is **positive semi-definite** if $\;Q(x) \geq 0$ for all $x \in V$.

- Q is **negative semi-definite** if $\;Q(x) \leq 0$ for all $x \in V$.

- Q is **indefinite** if it is neither positive semi-definite, nor negative semi-definite.

  This means that there are vectors u and v such that $\;Q(u) < 0 < Q(v)$.

- Q is **nondegenerate** if its polar form is nondegenerate.

  When V is finite dimensional, Q is nondegenerate if and only if its associated symmetric matrix is nonsingular.

3.2 N-ary quadratic forms (Quadratic forms on $\mathbb{K}^n$)

In this section we give the general expression of a quadratic form on a finite dimensional vector spaces. We shall see that such quadratic forms are either identically zero or can be described as homogeneous polynomials of degree 2 in finitely many indeterminates.

Proposition 3.2.1

A functional $Q : \mathbb{K}^n \longrightarrow \mathbb{K}$ is a quadratic form if and only if there exists a symmetric matrix $\mathbf{A}$ with coefficients in $\mathbb{K}$ and of order n, such that

$$Q(x) \;=\; x^T \mathbf{A} x\,, \qquad \forall\, x \in \mathbb{K}\,.$$

In other words, $Q(x)$ is either identically zero or is a 2-homogeneous polynomial of degree 2 of the n variables $x_1,\, x_2,\, \ldots,\, x_n \in \mathbb{K}$ that are the components of $x \in \mathbb{K}^n$.

Proof. Follows from Definition 3.1.1, Proposition 3.1.2 and Theorem 2.5.1.

Remark 3.2.2

A quadratic form on $\mathbb{K}^n$ is also called an *n-ary quadratic form* over $\mathbb{K}$.

In the cases of one, two, and three variables they are more precisely called *unary*, *binary*, and *ternary* and have respectively the following explicit forms :

$$\begin{aligned} Q(x) &= ax^2 && (n=1). \\ Q(x,y) &= ax^2 + bxy + cy^2 && (n=2). \\ Q(x,y,z) &= ax^2 + by^2 + cz^2 + dxy + exz + fyz && (n=3). \end{aligned}$$

Definition 3.2.3

The *rank* of a quadratic form defined on a finite dimensional vector space over $\mathbb{K}$ is the rank of its polar form. That is the rank of the symmetric matrix associated to the quadratic form (once a basis is fixed).

Definition 3.3.3 (Equivalent quadratic forms)

Let Q and R be two n-ary quadratic forms over $\mathbb{K}$. R is said to be *equivalent* to Q, if R can be transformed into Q by a non-singular transformation P of indeterminates;

$$R(X) \;=\; Q(PX)\,, \qquad \forall\, X \in \mathbb{K}^n\,.$$

This means that their respective corresponding symmetric matrices A and B (with respect to the canonical basis of $\mathbb{K}^n$) are congruent ($B = P^T A P$).

3.3 Reduction of Quadratic forms on finite dimensional real vector spaces

The expressions of a quadratic form can be simplified by choosing a suitabe basis. This can be seen as follows:

- either from the fact that a quadratic space $(\mathbb{R}^n\,,\,Q)$ also means a symmetric bilinear space $(\mathbb{R}^n\,,\,f)$; where f is the polar form of Q, which has a f-orthogonal basis according to Theorem ... .
  Indeed in a f-orthogonal basis $\mathscr{B} = (u_1\,,\ldots,\,u_n)$ of $\mathbb{R}^n$, we have for every
  $$\begin{aligned} X &= y_1u_1 + \ldots + y_nu_n \in \mathbb{R}^n\,, \\ Q(X) = f(X,X) &= \textstyle\sum_{i=1}^n \sum_{j=1}^n y_iy_jf\,(u_i\,,\,u_j) \\ &= \textstyle\sum_{i=1}^n f\,(u_i\,,\,u_i)\,y_i^2\,, \end{aligned}$$
- or from the fact that every real symmetric matrix admits for $\mathbb{R}^n$ equipped with the dot product, an orthonormal basis of eigenvectors according to Theorem 1.4.1.
  Indeed, letting $\mathbf{A}$ be the symmetric representative matrix of Q in the canonical basis of $\mathbb{R}^n$, and taking an orthonormal basis of eigenvectors $(v_1\,,\ldots,\,v_n)$ with respective real eigenvalues $\lambda_1\,,\ldots,\,\lambda_n$, we have for every
  $$\begin{aligned} X &= \xi_1v_1 + \ldots + \xi_nv_n \in \mathbb{R}^n\,, \\ Q(X) = X^T\mathbf{A}X &= \left(\textstyle\sum_{i=1}^n \xi_iv_i\right)\cdot\left[\mathbf{A}\left(\textstyle\sum_{j=1}^n \xi_jv_j\right)\right] \\ &= \left(\textstyle\sum_{i=1}^n \xi_iv_i\right)\cdot\left(\textstyle\sum_{j=1}^n \xi_j\mathbf{A}v_j\right) \\ &= \left(\textstyle\sum_{i=1}^n \xi_iv_i\right)\cdot\left(\textstyle\sum_{j=1}^n \xi_j\lambda_jv_j\right) \\ &= \textstyle\sum_{i=1}^n \lambda_i\xi_i^2\,. \end{aligned}$$

Consequently, we have

Reduction to diagonalized forms

Theorem 3.3.1 **(Diagonalisation of bilinear/Quadratic Forms)[16]**

Let n be a natural number and V be a n-dimensional euclidean space. Moreover let Q be a quadratic form on V and denote by f the polar form of Q. Then there exists for V an orthonormal basis $\mathscr{B} = (v_1\,,\ldots,\,v_n)$ for which $[f]_{\mathscr{B}} = \operatorname{diag}(\lambda_1\,,\ldots,\,\lambda_n)$ and so

$$f(X,Y) = \sum_{i=1}^n \lambda_ix_iy_i \qquad \text{for } X = x_1v_1+\ldots+x_nv_n \text{ and } Y = y_1v_1+\ldots+y_nv_n\,,$$

while

$$Q(X) \;=\; \sum_{i=1}^{n} \lambda_i x_i^2 \qquad \text{for} \;\; X = x_1 v_1 + \ldots + x_n v_n \,.$$

Quadratic forms can even be more reduced to a canonical form in which we have only the coefficients 1, -1 or 0 when $\mathbb{K} = \mathbb{R}$, and only the coefficients 1 or 0 when $\mathbb{K} = \mathbb{C}$.

Reduction to canonical forms

Definition 3.3.2

Let Q be a quadratic form, with polar form f, on a finite dimensional vector space V over $\mathbb{C}$, and suppose there is a basis $\mathcal{B}$ of V such that

$$[f]_{\mathcal{B}} \;=\; \begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$$

where I_r is the identity matrix of order r if $r \geq 1$, and vacuous if $r = 0$.
The matrix $\begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$ characterizes the *canonical form* of Q (over $\mathbb{C}$). This canonical form is described as

$$x_1^2 \,+\, \ldots \,+\, x_r^2 \,.$$

Definition 3.3.3

Let Q be a quadratic form, with polar form f, on a vector space V over $\mathbb{R}$, and suppose there is a basis $\mathcal{B}$ of V such that

$$[f]_{\mathcal{B}} \;=\; \begin{pmatrix} I_p & & \\ & -I_m & \\ & & 0 \end{pmatrix}$$

where for $k = p$ or $k = m$, I_k is the identity matrix of order k if $k \geq 1$, and vacuous if $k = 0$.
The matrix $\begin{pmatrix} I_p & & \\ & -I_m & \\ & & 0 \end{pmatrix}$ characterizes the *canonical form* of Q over $\mathbb{R}$. This canonical form is described as

$$x_1^2 \,+\, \ldots \,+\, x_p^2 \,-\, x_{p+1}^2 \,-\, \ldots \,-\, x_{p+m}^2 .$$

Theorem 3.3.4 (Canonical forms over $\mathbb{C}$)

Let V be a n-dimensional vector space over $\mathbb{C}$ and let Q be a quadratic form on V. Then Q has exactly one canonical form.

Proof Let r be the rank of Q and f be the polar form of Q. If $r = 0$, then the result is trivial. Suppose now that $r \geq 1$. Choose an f-orthogonal basis $(\tilde{v}_1, \ldots, \tilde{v}_n)$. Then this basis must contain exactly r anisotropic vectors. Reorder this basis as $(v_1, ..., v_n)$ such that

$$Q(v_1) \ldots Q(v_r) \neq 0 \qquad \text{and} \quad Q(v_{r+1}) = \ldots = Q(v_n) = 0\,.$$

For each $i = 1, \ldots, r$, choose a square root $\beta_i \in \mathbb{C}$ of $Q(v_i)$.
Consider now the new basis $\mathcal{B} = (w_1, ..., w_n)$ where

$$w_i = \begin{cases} \frac{1}{\beta_i} v_i & \text{if } 1 \leq i \leq r\,, \\ v_i & \text{if } r < i \leq n\,. \end{cases}$$

It follows that $[f]_{\mathcal{B}} = \begin{pmatrix} I_r & \\ & 0 \end{pmatrix}$.

Corollary 3.3.5

Two quadratic forms over $\mathbb{C}$ are equivalent if and only if they have the same canonical form.

Theorem 3.3.6 (Sylvester's law of inertia)

Let n be a natural number and let V be a n-dimensional vector space over $\mathbb{R}$. Moreover let Q be a quadratic form on V.
Then Q has exactly one (real) canonical form.
More precisely, by letting D be a diagonal form which is equivalent to Q and denoting by p (resp. m) the number of positive (resp. negative) coefficients of the expression of D, we have :

(i) $s = p - m$ is an invariant of Q, called the *signature* of Q and is denoted by $\operatorname{sig} Q$;

(ii) $p = \frac{r+s}{2}$ et $m = \frac{r-s}{2}$.

(iii) Q is equivalent to the canonical quadratic form defined for all $(x_1, \ldots, x_n) \in \mathbb{R}^n$ by

$$D_*(x_1, \ldots, x_n) = x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2\,,$$

(iv) a quadratic form Q_1 is equivalent to Q if and only if

$$\operatorname{rank} Q_1 = r \quad \text{and} \quad \operatorname{sig} Q_1 = s\,.$$

Proof.
Existence. Let r be the rank of Q and f be the polar form of Q. Note that $p+m=r$. If $r=0$, then the result is trivial. Suppose that $r \geq 1$. Choose an f-orthogonal basis $(\tilde{v}_1, \ldots, \tilde{v}_n)$. Then this basis must contain exactly r anisotropic vectors. Reorder this basis as $(v_1, ..., v_n)$ such that

$$Q(v_i) > 0 \quad \text{if} \quad 1 \leq i \leq p\,, \qquad Q(v_i) < 0 \quad \text{if} \quad p+1 \leq i \leq r$$
$$\text{and} \quad Q(v_i) = 0 \quad \text{if} \quad r+1 \leq i \leq n\,.$$

Now consider the new basis $\mathcal{B} = (w_1, ..., w_n)$ where

$$w_i \;=\; \begin{cases} \frac{1}{\sqrt{|Q(v_i)|}} v_i & \text{if} \quad 1 \leq i \leq r\,, \\ \\ v_i & \text{if} \quad r+1 \leq i \leq n\,. \end{cases}$$

It follows that $[f]_{\mathcal{B}} \;=\; \begin{pmatrix} I_p & & \\ & -I_m & \\ & & 0 \end{pmatrix}$.

Uniqueness. Suppose we have two bases $\mathcal{B}$ and $\mathcal{C}$ with

$$[f]_{\mathcal{B}} = \begin{pmatrix} I_p & & \\ & -I_m & \\ & & 0 \end{pmatrix}\,, \quad [f]_{\mathcal{C}} = \begin{pmatrix} I_{p'} & & \\ & -I_{m'} & \\ & & 0 \end{pmatrix}$$

By comparing the ranks, we know that $p+m = p'+m'$. It's therefore sufficient to prove that $p = p'$. To this end, define two subspaces of V as follows

$$U = \text{Span}\{v_1\,, ..., v_r\}, \qquad W = \text{Span}\{w_{p'+1}\,, ..., w_n\}.$$

If u is a non-zero vector of U, then we have $u = x_1 v_1 + ... + x_r v_r$, and hence

$$Q(u) = x_1^2 + ... + x_r^2 \;>\; 0\,.$$

Similarly if $w \in W$, then $w = y_{p'+1} w_{p'+1} + ... + y_n w_n$, and

$$Q(w) = -y_{p'+1}^2 - ... - y_{p'+s'}^2 \;\leq\; 0\,.$$

It follows that $U \cap W = 0$. Therefore

$$U + W \;=\; U \oplus W \;\subset\; V\,.$$

From this we have

$$\dim U \;+\; \dim W \leqslant \dim V\,.$$

Hence

$$p + (n - p') \leqslant n$$

This implies $p \leqslant p'$. A similar argument (by considering $U = \text{Span}\{v_1, ..., v_{p'}\}$ and $W = \text{Span}\{w_{p+1}, ..., w_n\}$) shows that $p' \leqslant p$, so we have $p = p'$. $\square$

As a corollary, we have the following result.

Proposition 3.3.7

Let Q be a quadratic form of rank $r > 0$ and signature s in n indeterminates over $\mathbb{R}$. Let $x_1, \ldots, x_n$ be the indeterminates and let $p = (r+s)/2$. Then

(i) Q is positive (resp. negative) definite if and only if $r = s = n$ (resp. $r = -s = n$); that is, Q is equivalent to the quadratic form defined by

$$D(x_1, \ldots, x_n) \;=\; x_1^2 + \ldots + x_n^2\,, \qquad \forall\,(x_1, \ldots, x_n) \in \mathbb{R}^n\,;$$

(resp. $-D$).

(ii) Q is positive (resp. negative) semi-definite if and only if $r = s$ (resp. r=-s); that is, Q is equivalent to the quadratic form defined by

$$D(x_1, \ldots, x_n) \;=\; x_1^2 + \ldots + x_r^2\,, \qquad \forall\,(x_1, \ldots, x_n) \in \mathbb{R}^n\,;$$

(resp. $-D$).

(iii) Q is indefinite if and only if $s < r$; that is, Q is equivalent to the quadratic form defined by

$$D(x_1, \ldots, x_n) \;=\; x_1^2 + \ldots + x_p^2 - x_{p+1}^2 - \ldots - x_r^2\,, \qquad \forall\,(x_1, \ldots, x_n) \in \mathbb{R}^n\,;$$

where $0 < p < r$.

3.4 Quadratic forms on Hilbert spaces

In an infinite dimensional real Hilbert space, if we have a quadratic form Q of which polar form f is is bounded, then by Theorem ..., there exists a bounded linear symmetric operator A such that

$$f(x, y) \;=\; \langle x, Ay \rangle\,, \qquad \forall x,\, y \in H\,,$$

and so

$$Q(x) \;=\; \langle x, Ax \rangle\,, \qquad \forall x \in H\,.$$

If moreover A is compact, and H is separable, then we can apply the spectral decomposition Theorem ... and get immediately a generalization of what is well-known in Euclidean spaces.

Theorem 3.4.1

Suppose A is a compact symmetric operator on a separable real Hilbert space H. There exists an at most countable orthonormal system

$$\big(\varphi_k\big)_{k \in \Gamma} \;\text{ where }\; \Gamma \subset \mathbb{N}\,,$$

of eigenvectors of A, with corresponding eigenvalues $\big(\lambda_k\big)_{k\in\Gamma}$ (converging to 0 if A is not of finite rank) such that for all $x \in H$,

$$Ax = \sum_{k\in\Gamma} \langle x, \varphi_k\rangle \varphi_k \,.$$

Hence, the corresponding quadratic form is defined by:

$$\begin{aligned} Q(x) = \langle x,\, Ax\rangle &= \textstyle\sum_{k\in\Gamma} \lambda_k \langle x, \varphi_k\rangle\langle x, \varphi_k\rangle \\ &= \textstyle\sum_{k\in\Gamma} \lambda_k |\langle x, \varphi_k\rangle|^2 \\ &= \textstyle\sum_{k\in\Gamma} \lambda_k x_k^2 \,, \end{aligned}$$

that is

$$Q(x) \;=\; \sum_{k\in\Gamma} \lambda_k x_k^2 \,, \qquad \text{where } \; x_k = \langle x, \varphi_k\rangle \;\text{ for all } x \in H \,.$$

In this case the study of the sign of Q reduces to the study of the signs of the eigenvalues $\big(\lambda_k\big)_{k\in\Gamma}$.

Besides there is another interesting case that involves unbounded operators !

Let H be a real Hilbert space with inner product denoted by

$\langle x,\, y\rangle$ for $x,\, y \in H$, and corresponding norm denoted by $||x||$ for $x \in H$.

Let V be another real Hilbert space with inner product denoted by $\langle x|y\rangle$ for $x,\, y \in V$, and corresponding norm denoted by $|x|$ for $x \in V$.
Suppose that:

1. V satisfies $V \subset H$ as a vector space and we have a continuous injection

   $$(V, |\cdot|) \;\hookrightarrow\; (H, ||\cdot||)$$

   in the sense that there exists a positive constant c such that

   $$|v| \;\le\; c||v|| \,, \qquad \forall v \in V \,.$$

2. V is dense in H. And it follows that $H^* \hookrightarrow V^*$.

Under these hypothesis, by Riesz representation Theorem, we have

$$V \;\hookrightarrow\; H \simeq H^* \;\hookrightarrow\; V^* \,,$$

and we call $\big(V,\, H,\, V^*\big)$ a *Gelfand triple.*

Theorem 3.4.2 [28]

Let $\big(V,\, H,\, V^*\big)$ be a Gelfand triple where H is separable and of infinite dimension. Let

$$a = a(\ ,\ )\ :\ V \times V \ \longrightarrow\ \mathbb{R}\ ; \quad (x, y) \mapsto a(x, y)\,,$$

be a symmetric, positive definite, and bounded.
Then :

1. there exists a bounded linear operator $\tilde{A}\,:\ V \to V^*$ and a linear (unbounded) operator $A\,:\ D(A) \subset H \to H$ defined respectively by

$$\Big\langle \tilde{A}x,\, y \Big\rangle_{V^*,\,V} \ =\ a(x, y)\,, \qquad \forall x,\, y \in H\,,$$

   and

$$A \ =\ \tilde{A}\Big|_{D(A)}\,, \quad \text{with} \quad D(A) = \Big\{ x \in V\ :\ \tilde{A}x \in H \Big\}\ .$$

2. The operator A is sefl-adjoint in H and bijective. And so $A^{-1}\,:\ H \to H$ is defined and self-adjoint.

3. If furthermore the continuous injection $V \hookrightarrow H$ is compact, then A^{-1} is a self-adjoint compact operator and the eigenvalue problem

$$v \in V \setminus \{0\}\,, \quad a(v, w) \ =\ \mu\langle v, w\rangle \quad \text{for every } w \in V\,,$$

   that amounts to

$$Av \ =\ \mu v \ \Longleftrightarrow\ v = \mu A^{-1} v\,,$$

   has a countably many positive eigenvalues $(\mu_k)_{k\in\mathbb{N}}$ with finite multiplicities such that

$$0 \ <\ \mu_1 \ <\ \mu_2 \ <\ \ldots \ \ \to\ +\infty\,,$$

   and the associated normalized system of eigenvectors forms a Hilbertian basis of H.
   Therefore, it is easy to see that

$$\mu_1 \ =\ \min\left\{ \frac{a(v, v)}{|v|^2}\ :\ \ v \in V \setminus \{0\}\,. \right\}$$

   The ratio

$$R(v) \ =\ \frac{a(v, v)}{|v|^2}$$

   is called the *Rayleigh quotient* and the previous identity (satisfied by μ_1) is called the *Rayleigh principle.*

Chapter 4

Applications

4.1 Quadratic forms and Unconstrained Optimization

Proposition 4.1.1 [23],[19]

Let H be a real Hilbert, Ω be a nonempty open set of H and $f : \Omega \to \mathbb{R}$ be a function. Let $x_0 \in \Omega$.

- If f is differentiable at x_0, then the derivative (in the sense of Fréchet) of f at x_0 is a bounded linear functional on H and so there exists a unique vector denoted by $\nabla f(x_0)$ and called the *gradient* of f at x_0 such that
$$f'(x_0)(h) \;=\; \langle \nabla f(x_0),\, h\rangle\,, \quad \forall h \in H\,.$$
- If f is of class $\mathcal{C}^2$, then the second order derivative (in the sense of Fréchet) of f at x_0 is a symmetric bounded bilinear form on H and so there exists a unique bounded symmetric operator denoted by $\mathcal{H}_f(x_0)$ and called the *Hessian* of f at x_0 such that
$$f''(x_0)(u,v) \;=\; \langle u,\, \mathcal{H}_f(x_0)v\rangle\,, \quad \forall u,\, v \in H\,.$$

Theorem 4.1.2 (Optimality Necessary Condition)[3]

Let Ω be a nonempty open set in $\mathbb{R}^n$, let f be a real-valued function defined on Ω and suppose that $x_0 \in \Omega$ is a local minimizer.

1. If f has first order partial derivatives at x_0 , then
$$\frac{\partial f}{\partial x_i}(x_0) \;=\; 0\,, \quad \text{for all } i = 1,\, 2,\, \ldots,\, n.$$
In particular, if f is differentiable at x_0, then x_0 is a critical point of f; that is,
$$\nabla f(x_0) \;=\; \left(\frac{\partial f}{\partial x_1}(x_0),\, \ldots,\, \frac{\partial f}{\partial x_n}(x_0)\right) \;=\; 0_{\mathbb{R}^n}\,.$$

2. If f is of class $\mathcal{C}^2$ on a neighbourhood of x_0, then not only we have $\nabla f(x_0) = 0$, but we also have that the Hessian of f at x_0 ;

$$\mathcal{H}_f(x_0) \;=\; \left(\frac{\partial^2 f(x)}{\partial x_i \partial x_j}\right)_{1\le i,j\le n}$$

is positive semi-definite.

Theorem 4.1.3 (Optimality Sufficient Condition)[3],[9]

Let Ω be a nonempty open set in $\mathbb{R}^n$, let f be a real-valued function differentiable on Ω and let $x_0 \in \Omega$ be a critical point of f.

1. If f is of class $\mathcal{C}^2$ on a neighbourhood of x_0 and $\mathcal{H}_f(x_0)$ is positive definite, then x_0 is a strict local minimizer.

2. If Ω is convex, f is of class $\mathcal{C}^2$ on Ω, and $\mathcal{H}_f(x)$ is positive semi-definite for all $x \in \Omega$, then x_0 is a (global) minimizer of f.

   If $\mathcal{H}_f(x)$ is positive definite for all $x \in \Omega$, then x_0 is a strict (global) minimizer of f.

Proof.
The proof follows from the Taylor expansion of f at x_0 with integral remainder. In fact for every $x \in \Omega$ such that the segment joining x_0 and x lies in an open set on which f is twice continuously differentiable, we have by setting $h = x - x_0$

$$\begin{aligned} f(x) \;=\; f(x_0+h) \;&=\; f(x_0) \,+\, \langle \nabla f(x_0), h\rangle \,+\, \textstyle\int_0^1 \langle h,\, \mathcal{H}(x_0+th)h\rangle\, dt \\ &=\; f(x_0) \,+\, \textstyle\int_0^1 \langle h,\, \mathcal{H}(x_0+th)h\rangle\, dt\,. \end{aligned}$$

The reader interested in the infinite dimensional version of Theorem 4.1.2 and Theorem 4.1.3 is refered to [25],[23],[9].

A straightforward application of these theorems yields

Example 4.1.4 (Quadratic Optimization)

Let n be a natural number and A be a real symmetric matrix of order n. Consider the elements of $\mathbb{R}^n$ as column vectors and let $b \in \mathbb{R}^n$. Define the function $f \,:\, \mathbb{R}^n \to \mathbb{R}$ by

$$f(x) \;:=\; \frac{1}{2}\langle x,\, Ax\rangle \,+\, \langle b,\, x\rangle \;=\; \frac{1}{2}x^T Ax \,+\, b^T x\,.$$

Then

1. f has a minimum if and only if the equation $Ax + b = 0_{\mathbb{R}^n}$ has a solution and A is positive semi-definite.

2. If A is positive definite, then f has a strict global minimum.

3. Any local minimum of f is a global minimum.

4.2 Optimization of Convex Functions

Definitions 4.2.1

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **convex** if for all x, $y \in \mathbb{R}^n$,

$$f(\lambda x + (\lambda - 1)y) \leq \lambda f(x) + (\lambda - 1)f(y)\,, \quad \forall \lambda \in [0, 1].$$

f is called **strictly convex** if this inequality is strict for all $x \neq y$ and $\lambda \in (0, 1)$.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is said to be **concave** if its opposite $-f$ is convex. That is:

$$f(\lambda x + (\lambda - 1)y) \geq \lambda f(x) + (\lambda - 1)f(y) \quad \forall x, y \in \mathbb{R}^n, \quad \forall \lambda \in [0, 1].$$

f is called strictly concave if the last inequality is strict for all $x \neq y$ and $\lambda \in (0, 1)$.

Characterization of convex functions

Proposition 4.2.2 (Slope inequality)

Let I be a nonempty interval of $\mathbb{R}$ and $\varphi : I \longrightarrow \mathbb{R}$.

(i) for all r, s, $t \in I$ such that $r < s < t$, we have

$$\frac{\varphi(s) - \varphi(r)}{s - r} \leq \frac{\varphi(t) - \varphi(r)}{t - r} \leq \frac{\varphi(t) - \varphi(s)}{t - s}$$

(and vice-versa).

(ii) In particular, given real numbers $a < b$ in I and $\delta > 0$ such that $2\delta < (b - a)$, we have

$$\frac{\varphi(a + \delta) - \varphi(a)}{\delta} \leq \frac{\varphi(t) - \varphi(s)}{t - s} \leq \frac{\varphi(b) - \varphi(b - \delta)}{\delta} \tag{4.2.1}$$

$$\text{for} \quad a + \delta \leq s < t \leq b - \delta\,,$$

showing that φ is locally Lipschitz.

Likewsise, if φ is strictely convex, then for all r, s, $t \in I$ such that $r < s < t$, we have

$$\frac{\varphi(s) - \varphi(r)}{s - r} < \frac{\varphi(t) - \varphi(r)}{t - r} < \frac{\varphi(t) - \varphi(s)}{t - s}$$

(and vice-versa).

Proposition 4.2.3

Let Ω be a nonempty open convex set of $\mathbb{R}^n$ and $f : \Omega \longrightarrow \mathbb{R}$ be a function which is (Fréchet) differentiable. Then the following three assertions are equivalent:

(i) f is convex.

(ii) $\langle \nabla f(y) - \nabla f(x),\, y - x \rangle \geq 0$ for all $x,\, y \in \Omega$; (i.e., ∇f is monotone).

(iii) $\langle \nabla f(x),\, y - x \rangle \;\leq\; f(y) - f(x)$ for all $x,\, y \in \Omega$.

A function f which is twice differentiable on Ω is convex if and only if for each $x \in \Omega$ the Hessian matrix $\mathcal{H}(x)$ is positive semidefinite.

Similarly, the following are equivalent :

a) f is strictly convex.

b) $\langle \nabla f(y) - \nabla f(x),\, y - x \rangle \;>\; 0 \quad$ for all $x,\, y \in \Omega$ with $x \neq y$.

c) $\langle \nabla f(x),\, y - x \rangle \;>\; f(y) - f(x)$ for all $x,\, y \in \Omega$ with $x \neq y$.

Furthermore if f is twice differentiable and its hessian at every $x \in \Omega$ positive definite, then f is strictly convex.

Consequently, given a real symmetric matrix A of order n, the function $f : \mathbb{R}^n \to \mathbb{R}$ defined by

$$f(x) \;:=\; \frac{1}{2}\langle x,\, Ax \rangle \;+\; \langle b,\, x \rangle \;=\; \frac{1}{2} x^T A x \;+\; b^T x$$

is :

1. convex if and only if A is positive semi-difinite.
2. strictly convex if and only if A is positive difinite.
3. concave if and only if A is negative semi-difinite..
4. strictly concave if and only if A is negative semi-difinite..
5. is neither convex nor concave if and only if A is indefinite.

Furthermore Proposition 4.2.3 shows that the second part of Theorem 4.1.3 is just about the Optimization of Convex Functions of class $\mathcal{C}^2$.

Example 4.2.4 (Linear Regression in Statistics)

In the linear regression problem, n points (x_1, y_1), (x_2, y_2), $\ldots$, (x_n, y_n) are given in the xy-plane and it is required to "fit" a straight line with equation $y = ax + b$ in such a way that the sum of the squares of the vertical distances of the given points from the line is minimized. That is, (a, b) is to be chosen in $\mathbb{R}^2$ so that the binary-quadratic form

$$Q(a, b) \;=\; \sum_{i=1}^{n} (ax_i + b - y_i)^2$$

is minimized.

When the x_i's are not all equal, such a line exits and is called the *regresion line.*

We obtain Its corresponding coefficients as

$$b \;=\; \overline{y} - a\overline{x}\;, \qquad a \;=\; \frac{n\overline{x}\,\overline{y} - \sum_{i=1}^{n} x_i y_i}{n\overline{x}^2 - \sum_{i=1}^{n} x_i^2}\;;$$

where

$$\overline{x} \;=\; \frac{1}{n}\sum_{i=1}^{n} x_i\;, \qquad \text{and} \quad \overline{y} \;=\; \frac{1}{n}\sum_{i=1}^{n} y_i\;.$$

Example 4.2.5 (Mean-Square Approximation)

(Best Mean-Square Approximation by Orthogonal Functions, Fourier coefficients)

Let φ_1, φ_1, $\ldots$, φ_n be continuous and not identically zero on a compact interval $[a, b]$ where $a < b$ and such that

$$\int_a^b \varphi_i(x)\varphi_j(x)\,dx \;=\; 0 \qquad \text{if } i \neq j\,.$$

Given a piecewise continuous function on $[a, b]$, we can see by Theorem ... that the values of c_1, c_2, $\ldots$, c_n that minimize the n-ary quadratic form

$$Q(c_1\,,\, c_2\,,\, \ldots,\, c_n) \;=\; \int_a^b \left| f(x) \;-\; \sum_{i=1}^{n} c_i\varphi_i(x) \right|^2 dx\;,$$

are given by the formulas

$$c_i \;=\; \frac{\langle \varphi_i\,,\, f\rangle}{\langle \varphi_i\,,\, \varphi_i\rangle}\;, \qquad i = 1,\, \ldots,\, n\,,$$

where

$$\langle \varphi,\, \psi\rangle \;=\; \int_a^b \varphi(x)\psi(x)\,dx\,.$$

Example 4.2.6 (Lax-Milgram Theorem)

Let H be a real Hilbert space and

$$a : H \times H \to \mathbb{R}$$

be a bounded bilinear form that is coercive so that there exist positive constants M and c such that

$$|a(x,y)| \;\leq\; M||x||\,||y|| \qquad \forall\, x,\, y \in H\,,$$

and

$$a(x,x) \;\geq\; c||x||^2 \qquad \forall\, x \in H\,.$$

1. Suppose moreover that a is symmetric.

   i) Then it is not hard to see that the mapping

   $$(x,y) \;\longmapsto\; \langle x,y\rangle_a \;=\; a(x,y)$$

   is an inner product on H.

   ii) Moreover the new norm

   $$||\cdot||_a \;:\; x \;\longmapsto\; ||x||_a \;=\; \sqrt{a(u,u)}$$

   induced by the inner product $\langle\,,\,\rangle_a$ is equivalent to the norm $||\cdot||$ of H.

   iii) Therefore,

   for any bounded linear functional f of H,

   $$f \;:\; (H,\, ||\cdot||) \;\longrightarrow\; \mathbb{R}\,;$$

   that is $f \in H^*$, $\big(f$ is also a bounded linear functional of $(H,\, \langle\,,\,\rangle_a)$, and thus$\big)$ there exists a unique solution $v \in H$ of the variational equation

   $$a(x,v) \;=\; f(x) \qquad \text{pour tout } x \in H$$

   with

   $$||v|| \;\leq\; \frac{1}{c}||f||_{H^*}\,.$$

   The same conclusion can be achieved by minimizing the coercive and continuous functional $F : H \to \mathbb{R}$ defined by

   $$F(x) \;=\; \frac{1}{2}a(x,x) \;-\; f(x)\,, \qquad x \in H\,;$$

where f is given in H^*. This functional F is differentiable and its derivative at each $x_0 \in H$ is defined by

$$\begin{aligned} F'(x_0)(h) &= \tfrac{1}{2}\big(a(h,x_0) + a(x_0,h)\big) - f(h) \\ &= a(h,x_0) - f(h)\,, \quad \forall h \in H\,. \end{aligned}$$

The unique zero of F' is our solution.

2. General case (a needs not be symmetric).
In this case it suffices to see that the bounded linear operator A associated to the bounded bilinear form a by theorem ... and satisfying

$$a(x,y) = \langle x,\, Ay\rangle\,, \qquad \forall\, x,\, y \in H\,,$$

is bijective. This is achieved by the following steps:

i) $c||x|| \le ||Ax|| \le M||x||$ pour tout $x \in H$,
which implies imediately that A is one-to-one.

ii) $(\operatorname{Im} A)^{\perp} = \{0\}$ and $\operatorname{Im} A$ is closed.
This implies that A is onto.
(The closedness of $\operatorname{Im} A$ follows from the fact that for any convergent sequence $(y_n)_n$ of elements of $\operatorname{Im} A$ with $y_n = Ax_n$, $x_n \in H$, we have

$$||x_n - x_m|| \le \frac{1}{c}||Ax_n - Ax_m|| = ||y_n - y_m|| \qquad \text{for all } n,\, m$$

showing that $(x_n)_n$ is a cauchy sequence of the Hilbert space H and must converge. We are done since A is continuous).

iii) Therefore given a bounded linear functional f on H, we have

$$a(x,v) = f(x) \;\; \forall x \in H \quad \Longleftrightarrow \quad \langle x,\, Av\rangle = f(x) \;\; \forall x \in H\,.$$

Denoting by $z \in H$, the Riesz representation of f, we have:

$$Av = z$$

and so

$$v = A^{-1}z$$

with

$$||v|| \le ||A^{-1}||\,||z|| = ||A^{-1}||\,||f|| \le \frac{1}{c}||f||\,.$$

For concrete examples concerning the use of Lax-Milgram Theorem or Gelfand triple to solve Dirichlet problems, the interested reader is refered to [6] and [19].

Bibliography

[1] Lang, S. : Calculus of several variables. 2nd Edition. Springer New-York. 1987.

[2] Jost, J. : Postmodern Analysis. Springer Berlin. 1998.

[3] Berkovitz, L.D. : Convexity and Optimization in $\mathbb{R}^n$. John Wiley & Sons, Inc. 2001.

[4] Pfister, A. : Quadratic Forms with Applications to Algebraic Geometry and Topology. Cambridge University Press. 1995.

[5] Gregory, J. : Quadratic Form Theory and Differential Equations. Mathematics in Sciences and Engineering 152. Academic Press 1980.

[6] Lax, P. : Functional Analysis. Wiley-Interscience 2002.

[7] O'Meara, O.T. : Introduction to Quadratic Forms. 3rd Ed. Springer Berlin 1973.

[8] Simon, B. : *Hamiltonians defined as quadratic forms.* Comm. Math. Phys. 21. (1971), 192-210.

[9] Troutman, J.L. : Variational Calculus and Optimal Control. Optimization with elementary convexity. 2nd Ed. Springer 1996

[10] Simon, B. : *A canonical decomposition of quadratic forms with applications to monotone operators.* Journal of Functional Analysis 28. (1978) 337-385.

[11] Kuhnel, W. : Differential Geometry -Curves-Surfaces-Manifolds. 2nd Ed. AMS 2006.

[12] Golubitsky, M. and Marsden, J. : *The Morse lemma in infinite dimensions in Singularity Theory.* SIAM J. Math Anal. Vol. 14(6). (1983). 1037-1044.

[13] Salako, R. : Foundation of Morse Theory. AUST MSc Dissertation. Abuja 2011.

[14] Serre, J-P. : Lie Algebras and Lie Groups. LNM 1500. 2nd Ed. Springer Berlin 1992.

[15] Sabinin, L.V. : Mirror Geometry of Lie Algebras, Lie Groups and Homogeneous Spaces. MA 573. Kluwer Academic Publishers

[16] Bowers, J. : Matices and quadratic forms. 2001

[17] Muscat, J. : An introduction to Metric spaces, Hilbert spaces and Banach algebras. Springer 2014.

[18] Bressan, A. : Lecture Notes on Functional Analysis with Applications to Partial Differential Equations. GSM 143. AMS 2013.

[19] Brezis, H. : Functional Analysis, Sobolev Spaces and Partial Differential Equations. Springer New-York 2011.

[20] Fabian, M., Habala, P., Montesinos, V. and Zisler, V. : Banach Space Theory. The basis for linear and nonlinear analysis. Springer 2011.

[21] Chidume, C.E. : Applicable Functional Analysis. Trieste 2006.

[22] Yosida, K. : Functional Analysis. Springer Berlin. 1980.

[23] Amann, H. and Escher, J. : Analysis II. Birkhauser 1999.

[24] Hansen, V.L. : Functional Analysis. Entering Hilbert space. World Scientific 2006.

[25] Blanchard, P. and Bruning, E. : Variational Methods in Mathematical Physics. A unified approach. Springer Verlag 1992.

[26] Anderson, T.N. and Taylor, J.B. : *An inequality for a sum of quadratic forms with applications to Probability Theory.* Linear Algebra and its Applications. Vol. 30 (1980). 93-99.

[27] Goldberg, G. : *On the numerical radius and its applications.* Linear Algebra 42. (1982) 263-284.

[28] Senba, T. and Suzuki, T. : Applied Analysis. Mathematical methods in Natural sciences. Imperial College Press 2004.

[29] Elduque, A. and Perez, J.M. : *Infinite dimensional Quadratic forms admitting composition.* Proc. of the AMS. 125(8). 1997. 2207-2216.

[30] Bellman, R. : Introduction to Matrix Analysis. 2nd Edition. SIAM Classics in Application 19. The rand cooperation 1997.

[31] Dal Maso, G. : An Introduction to Gamma-convergence. Birkhauser 1992.

[32] Daddy, A.J. : Spectral Theory for Compact Linear Operators and Applications. AUST MSc Dissertation. Abuja 2011.

Index