INF-SUP CONDITIONS FOR OPERATOR EQUATIONS

LONG CHEN

CONTENTS

We study the well-posedness (existence, uniqueness and stability) of the equation

$$Tu = f, \tag{1}$$

where T is a linear and bounded operator between two linear vector spaces. We give equivalent conditions on the existence and uniqueness of the solution and apply to variational problems to obtain the so-called inf-sup condition (also known as Babuška condition [1]). When the linear system is in the saddle point form, we derive another set of inf-sup conditions (known as Brezzi conditions [**?**]).

When linear spaces are finite dimensional and therefore T is a matrix, the existence and uniqueness of solution to (1) is equivalent to T is square and non-singular. The stability, i.e., the solution u depends continuously on f characterized by $\|u\| \leq C\|f\|$, is equivalent to $\sigma_{\min}(T) > 0$. The reason we study the operator equation (1) in infinite dimensional spaces is that in numerical methods, there is a sequence of finite dimensional matrix equations $T_h u_h = f_h$ and the stability should be uniform to the parameter h so that we can safely take the limit $h \to 0$ and obtain the order of convergence of $\|u - u_h\|$.

Date: Last update: January 13, 2022.

1. Preliminary from Functional Analysis

In this section we recall some basic facts in functional analysis, notably three theorems: Hahn-Banach Theorem, Closed Range Theorem, and Open Mapping Theorem. For detailed explanation and sketch of proofs, we refer to *Chapter: Minimal Functional Analysis for Computational Mathematicians*.

1.1. **Function Analysis.** Calculus is on the relation (differentiation and integration) of a function f and the independent variable x, e.g. $\lim_{x\to x_0} f(x)$. Real analysis studies the space consisting of functions such as convergence of a sequence of functions, e.g. $\lim_{n\to\infty} f_n$. Functional analysis moves one level higher and focus on the operators between these function spaces, e.g. $\lim_{n\to\infty} T_n(f)$. One important example is the functional which can be defined as a function of functions.

Linear algebra, which can be also phrased as matrix analysis, is studying the operators between finite dimensional spaces. Spaces studied in the functional analysis, is in general infinitely dimensional and thus an operator can be thought as a infinite dimensional matrix.

Simply speaking, functional analysis is the combination of linear algebra and analysis in infinite dimensional linear spaces. To transfer results in calculus and linear algebra to functional analysis, additional conditions should be imposed on the space and/or linear operators since the dimension of function spaces are in general infinite which brings a lot of room for many non-intuitive facts. Among many others, the following three properties are extremely important: *closeness or completeness, continuity (boundedness), and compactness*.

Since we will mainly deal with spaces of functions, we use notation u, v or f, g as elements in the linear space of functions and reserve x, y for points in $\mathbb{R}^n$. The most interesting object is $u = u(x)$, i.e., spaces consisting of functions u of x. The domain for x is usually denoted by Ω. Note that as a subset in $\mathbb{R}^n$, properties of Ω (closed or open, compact, or bounded) will also affect (but not directly related to) properties of the function space defined on Ω.

1.2. **Spaces.** A complete normed space will be called a *Banach space*. A complete inner product space will be called a *Hilbert space*. Completeness means every Cauchy sequence will have a limit and the limit is in the space. Completeness is a nice property so that we can safely take the limit.

Functional analysis is studying $\mathscr{L}(U, V)$: the linear space consisting of all linear operators between two vector spaces U and V. When U and V are topological vector spaces (TVS), the subspace $\mathscr{B}(U, V) \subset \mathscr{L}(U, V)$ consists of all continuous linear operators. An operator T is continuous if the pre-image of any open set is open. An operator T is bounded if T maps bounded sets into bounded sets. When U and V are normed spaces, for a bounded operator, there exists a constant M s.t. $\|Tu\| \leq M\|u\|, \forall u \in U$. The smallest constant M is defined as the norm of T. For $T \in \mathscr{B}(U, V)$, $\|T\| = \sup_{u\in U, \|u\|=1} \|Tu\|$. With such norm, the space $\mathscr{B}(U, V)$ becomes a normed vector space.

One can easily show that, for a linear operator, T is continuous iff T is bounded. Indeed by the translation invariance and the linearity of T, it suffices to prove such result at 0 for which the proof is straightforward by definition.

An important and special example is $V = \mathbb{R}$. The space $\mathscr{L}(U, \mathbb{R})$ is called the (algebraic) dual space of U and denoted by U^*. For an operator $T \in \mathscr{L}(U, V)$, it induces an operator $T^* \in \mathscr{L}(V^*, U^*)$ by $\langle T^* f, u\rangle = \langle f, Tu\rangle$, where $\langle \cdot, \cdot\rangle$ is the duality pair, and T^* is called the dual of T. The continuous linear functional of a normed space V will be denoted by V', i.e., $V' = \mathscr{B}(V, \mathbb{R})$ and the continuous dual T' is defined as T^*. The algebraic dual space only uses the linear structure while the continuous dual space needs a topology (to define the continuity). Note that since $\mathbb{R}$ is a Banach space, V' is a Banach space no matter V is or not. In general, when the target space is complete, we can define the limit of a sequence of operators and show that the space of bounded operators is also complete. Namely if U is a normed linear space and V is a Banach space, then $\mathscr{B}(U,V)$ is Banach.

Through the duality pair $\langle u, f\rangle$, we can identify an element $u \in V$ as an element in V'' and obtain an embedding $V \subset V''$. If $V'' = V$, then space V is called reflective. In general V' is a space different to V except for Hilbert space for which Reisz representation theorem shows V is isometric to V'.

Question: Why are the dual space and the dual operator so important?

(1) For an inner product space, the inner product structure is quite useful. For normed linear spaces, the duality pair $\langle\cdot,\cdot\rangle : V' \times V \to \mathbb{R}$ can play the (partial) role of the inner product.
(2) Since U' and V' are Banach spaces, T' is "nicer" than T. Many theorems are available for continuous linear operators between Banach spaces.

1.3. **Hahn-Banach Theorem.** A subspace S of a linear space V is a subset such that itself is a linear space with the addition and the scalar product defined for V. For a normed TVS, a closed subspace means the subspace is also closed under the topology, i.e., for every convergent sequence, the limit also lies in the subspace.

Theorem 1.1 (Hahn-Banach Extension)**.** *Let V be a normed linear space and $S \subset V$ a subspace. For any $f \in S' = \mathscr{B}(S,\mathbb{R})$ it can be extended to $f \in V' = \mathscr{B}(V,\mathbb{R})$ with preservation of norms.*

For a continuous linear functional defined on a subspace, the natural extension by density can extend the domain of the operator to the closure of S. So we can take the closure of S and consider closed subspaces only. The following corollary says that we can find a functional to separate a point with a closed subspace.

Corollary 1.2. *Let V be a normed linear space and $S \subset V$ a closed subspace. Let $v \in V$ but $v \notin S$. Then there exists a $f \in V'$ such that $f(S) = 0$ and $f(v) = 1$ and $\|f\| = \operatorname{dist}^{-1}(v,S)$.*

Proof. Consider the subspace $S_v = \operatorname{span}(S, v)$. For any $u \in S_v$, $u = u_s + \lambda v$ with $u_s \in S, \lambda \in \mathbb{R}$, we define $f(u) = \lambda$ and use Hahn-Banach theorem to extend the domain of f to V. Then $f(S) = 0$ and $f(v) = 1$ and it is not hard to prove the norm of $\|f\| = 1/d$. $\square$

The corollary is obvious in an inner product space. We can use the vector $\tilde{f} = v - \operatorname{Proj}_S v$ which is orthogonal to S and scale $\tilde{f}$ with the distance such that $f(v) = 1$. The extension of f is through the inner product. Thanks to the Hahn-Banach theorem, we can prove it without the inner product structure.

Another corollary resembles the Reisz representation theorem.

Corollary 1.3. *Let V be a normed linear space. For any $v \in V$, there exists a $f \in V'$ such that $f(v) = \|v\|^2$ and $\|f\| = \|v\|$.*

Proof. For a Hilbert space, we simply chose $f_v = v$ and for a norm space, we can apply Corollary 1.2 to $S = \{0\}$ and rescale the obtained functional. $\square$

Exercise 1.4. Prove $\|T\| = \|T'\|$ for $T \in \mathscr{B}(U,V)$ and U, V are normed linear spaces.

The norm structure in Hahn-Banach theorem is not necessary. It can be relaxed to a sub-linear functional and the preservation of norm can be relaxed to the preservation of an inequality.

Theorem 1.5 (Generalized Hahn-Banach Theorem)**.** *Let V be a linear space and S be a subspace of V. Let $p : V \to \mathbb{R}$ be a sub-linear functional. For any linear functional $f \in S'$ satisfying $f(v) \le p(v)$ for all $v \in S$, it can be extended to V' and still satisfies the same inequality for $v \in V$.*

Proof. We can modify the proof of Corollary 1.2 to construct an extension from S to $S+v$ as follows: for $u = u_s + \lambda v$, we define $f(v) = f(u_s) + c\lambda$. The constant c is carefully chosen to satisfy the constraint $f(v) \le p(v)$. We keep doing this until 'no point left'. Notice that the space is in general infinite dimensional (not even countable). To make the arguments more rigorous, we collect all extensions $\mathcal{K}$ and introduce a partially ordering on $\mathcal{K}$. Then we chose the maximal element in $\mathcal{K}$ and show the domain of this element is V by contradiction. Can we attain the maximal element? The answer is yes, if we can apply the Zorn's lemma on the axiom of choice. □

1.4. **Closed Range Theorem.** For an operator $T : U \to V$, denoted by $R(T) \subset V$ the range of T and $N(T) \subset U$ the null space of T. For a matrix $A_{m\times n}$ treating as a linear operator from $\mathbb{R}^n$ to $\mathbb{R}^m$, there are four fundamental subspaces $R(A), N(A^{\intercal}) \subset \mathbb{R}^m, R(A^{\intercal}), N(A) \subset \mathbb{R}^n$ and the following relation (named *the fundamental theorem of linear algebra* by G. Strang [3]).

Theorem 1.6 (The fundamental theorem of linear algebra [3])**.**

$$R(A) \oplus^{\perp} N(A^{\intercal}) = \mathbb{R}^m, \tag{2}$$

$$R(A^{\intercal}) \oplus^{\perp} N(A) = \mathbb{R}^n. \tag{3}$$

In words, the range space is the orthogonal complement of the null space. Proof is simply by definition of the adjoint operator

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

It is illustrated in the following figure.

We shall try to generalize (2)-(3) to operators $T \in \mathscr{B}(U,V)$ between normed/inner product spaces. For $T \in \mathscr{B}(U,V)$, due to the continuity of T, the *null space* $N(T) := \{u \in U, Tu = 0\}$ is a closed subspace. For a subset S in a Hilbert space H, the *orthogonal complement* $S^{\perp} := \{u \in H, (u,v) = 0, \forall v \in S\}$ is a closed subspace. For Banach spaces, we do not have the inner product structure but can use the duality pair $\langle \cdot, \cdot \rangle : V' \times V \to \mathbb{R}$ to define an "orthogonal complement" in the dual space which is called *annihilator*. More specifically, for a subset S in a normed space V, the annihilator $S^{\circ} = \{f \in V', \langle f, v\rangle = 0, \forall v \in S\}$. Similarly for a subset $F \subset V'$, we define ${}^{\circ}F = \{v \in V, \langle f, v\rangle = 0, \forall f \in F\}$. Similar to the orthogonal complement, annihilators are closed subspaces.

For a subset (not necessarily a subspace) $S \subset V$, $S \subseteq S^{\perp\perp}$ if V is an inner product space or $S \subseteq {}^{\circ}(S^{\circ})$ if V is a normed space. The equality holds if and only if S is a closed subspace (which can be proved using Hahn-Banach theorem). The space $S^{\perp\perp}$ or ${}^{\circ}(S^{\circ})$ is the smallest closed subspace containing S.

The range $R(T)$ is not necessarily closed even T is continuous. As two closed subspaces, the relation $N(T') = R(T)^{\circ}$ can be easily proved by definition. But the relation $R(T) = {}^{\circ}N(T')$ may not hold since ${}^{\circ}N(T')$ is closed but $R(T)$ may not.

Theorem 1.7 (Closed Range Theorem)**.** *Let U and V be Banach spaces and let $T \in \mathscr{B}(U,V)$. Then the following conditions are equivalent*

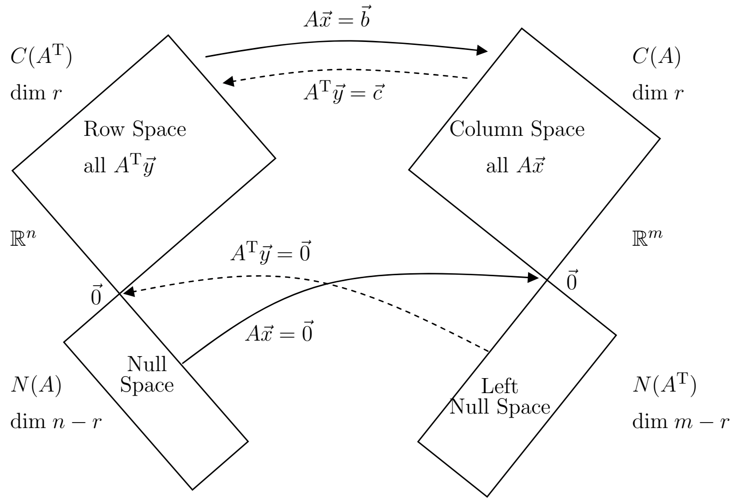

FIGURE 1. The fundamental theorem of linear algebra. Extract from G. Strang. *Linear Algebra and Its Applications* [4].

(1) $R(T)$ *is closed in* V.
(2) $R(T')$ *is closed in* U'.
(3) $R(T) = {}^\circ N(T')$
(4) $R(T') = N(T)^\circ$.

Proof. We give a proof from (1) to (3). The relation $R(T) = {}^\circ N(T')$ can be verified by definition. Suppose there exists $v \in {}^\circ N(T')$ but $v \notin R(T)$. Then by Corollary 1.2, there exists a $f \in V'$ s.t. $f(R(T)) = 0$ and $f(v) = 1$. The fact $\langle f, Tu\rangle = \langle T'f, u\rangle = 0, \forall u \in U$ implies $T'f \in N(T')$. The fact $v \in {}^\circ N(T')$ implies $f(v) = 0$ which contradicts with $f(v) = 1$. $\square$

Closeness is another nice property. An *operator is closed* if its graph is closed in the product space. More precisely, let $T : U \to V$ be a function and the graph of T is $G(T) = \{(u, Tu) : u \in U\} \in U \times V$. Then T is closed if its graph G is closed in $U \times V$ in the product topology. One can easily show a linear and continuous operator is closed. The definition of closed operators only uses the topology of the product space. A closed operator is not necessarily linear or continuous. When $T \in \mathscr{L}(U, V)$ and U, V are Banach spaces, these two properties are equivalent which is known as the closed graph theorem.

Theorem 1.8 (Closed Graph Theorem)**.** *Let* U *and* V *be Banach spaces and let* $T \in \mathscr{L}(U, V)$. *Then* T *is closed if and only if* T *is continuous.*

The requirement of U and V are Banach is necessary. Namely there exists a closed operators which is not continuous.

The range of a closed linear operator between Banach spaces (and thus continuous) is not necessarily closed. A simple example is the identity map $C([0, 1]) \to L^1([0, 1])$. With $\|\cdot\|_\infty$, $C([0, 1])$ is a Banach space and of course, the identity operator is continuous, and thus by the closed graph theorem, is closed. But $C([0, 1])$ is not closed with respect to the $\|\cdot\|_{L^1}$-norm. The sequence $\{\phi_n\}$ consists of hat function with height 1 and width $2/n$ will have $\|\phi_n\|_\infty = 1$ but $\|\phi_n\|_{L^1} = 1/n \to 0$. Later on we shall show to have a closed range, $\|Tu\| \geq c\|u\|$ should be bounded below.

Range is closed is stronger than the operator is closed. Compare the convergence in the definitions:

- Graph is closed: if $(u_n, Tu_n) \to (u, v)$, then $v = Tu$.
- Range is closed: if $Tu_n \to v$, then there exists a u such that $v = Tu$.

The difference is: in the second line, we do not know if u_n converges or not. But in the first line, we assume such limit exists.

Closedness is a relative concept depending on the topology. $C([0,1])$ is closed under the norm $\|\cdot\|_\infty$ but not closed w.r.t. $\|\cdot\|_{L^1}$-norm. In the closed graph theorem, the first component of the product topology of $U \times V$ will use the one for U. The sequence $\{Tu_n\}$ is Cauchy in V may not imply $\{u_n\}$ is Cauchy in U.

1.5. **Open Mapping Theorem.** The stability of the equation can be ensured by the open mapping theorem. An open map is a function between two topological spaces that maps open sets to open sets.

Theorem 1.9 (Open Mapping Theorem). *For $T \in \mathscr{B}(U, V)$ and both U and V are Banach spaces. If T is onto, then T is open.*

Usually the proof uses the Baire category theorem, and completeness of both U and V is essential. Again we skip the proof and consider a special case when T is bijective and use the closed graph theorem to give a short proof.

Theorem 1.10 (Banach Theorem). *For $T \in \mathscr{B}(U, V)$ and both U and V are Banach spaces. If T is into and onto, then T^{-1} exists and continuous.*

Proof. The graph of $G(T)$ and $G(T^{-1})$ are reflection of each other. Then T is continuous $\Rightarrow G(T)$ is closed $\iff G(T^{-1})$ is closed $\Rightarrow T^{-1}$ is continuous by closed graph theory. □

We then return to the case T is surjective. Let $S \subset V$ be a subspace. We can define a quotient space V/S using the linear structure only. When V is normed, we can define a quotient norm

$$\|u\|_{V/S} = \inf_{s \in S} \|u - s\|.$$

But this is only a semi-norm. For example, consider S be the subspace of all polynomials and Weierstrass approximation theorem says any continuous function can be approximated arbitrarily close by polynomials in the maximum norm. The polynomial subspace is not closed w.r.t. $\|\cdot\|_\infty$. It will be a norm if and only if S is a closed subspace. If V is a Banach space and $S \subset V$ is a closed subspace, then both S and V/S are Banach spaces.

When T is onto, the induced map

$$T : U/N(T) \to V$$

is bijective between two Banach spaces as $N(T)$ is a closed subspace. So we conclude $T^{-1} : V \to U/N(T)$ is continuous and thus $T : U/N(T) \to V$ is open. Now use relation of topology between $U/N(T)$ and U, we conclude the original T is also open.

2. INF-SUP CONDITION: BABUŠKA THEORY

The well-posedness of the operator equation $Tu = f$ consists of three questions: existence, uniqueness, and stability.

2.1. **Operator Equations.** For the uniqueness, a useful criterion to check is wether T is bounded below.

Lemma 2.1. *Let U and V be Banach spaces. For $T \in \mathscr{B}(U,V)$, the range $R(T)$ is closed and T is injective if and only if T is bounded below, i.e., there exists a positive constant c such that*

$$\|Tu\| \geq c\|u\|, \quad \textit{for all } u \in U. \tag{4}$$

Proof. **Sufficient**. If $Tu = 0$, inequality (4) implies $u = 0$, i.e., T is injective. Choosing a convergent sequence $\{Tu_k\}$, by (4), we know $\{u_k\}$ is also a Cauchy sequence and thus converges to some $u \in U$. The continuity of T shows that Tu_k converges to Tu and thus $R(T)$ is closed.

Necessary. When the range $R(T)$ is closed, as a closed subspace of a Banach space, it is also Banach. As T is injective, T^{-1} is well defined on $R(T)$. Apply Open Mapping Theorem to $T : U \to R(T)$, we conclude T^{-1} is continuous. Then

$$\|u\| = \|T^{-1}(Tu)\| \leq \|T^{-1}\|\|Tu\|$$

which implies (4) with constant $c = \|T^{-1}\|^{-1}$. □

When T is bounded below, we have $R(T)$ is closed and T is injective. As a closed subspace of a Banach space, $R(T)$ is also Banach, therefore

$$T : U \to R(T) = {}^{\circ}N(T')$$

is isomorphism (i.e. T and T^{-1} are linear and continuous).

A trivial answer to the existence of the solution to (1) is: if $f \in R(T)$, then it is solvable. When is it solvable for all $f \in V$? The answer is $V = R(T)$, i.e., T is surjective. A characterization can be obtained using the dual of T.

Lemma 2.2. *Let U and V be Banach spaces and let $T \in \mathscr{B}(U,V)$. Then T is surjective if and only if T' is injective and $R(T')$ is closed.*

Proof. **Sufficient**. By closed range theorem, $R(T)$ is also closed. Suppose $R(T) \neq V$, i.e., there exists a $v \in V$ but $v \notin R(T)$. By Hahn-Banach theorem, there exists a $f \in V'$ such that $f(R(T)) = 0$ and $f(v) = 1$. Then $T'f \in U'$ satisfies

$$\langle T'f, u\rangle = \langle f, Tu\rangle = 0, \quad \forall u \in U. \tag{5}$$

So $T'f = 0$ which implies $f = 0$ contradicts with the fact $f(v) = 1$.

Necessary. When T is surjective, then $R(T) = V$ is closed. By closed range theorem, so is $R(T')$. To prove T' is injective, we then show if $T'f = 0$, then $f = 0$. Indeed by (5), $\langle f, Tu\rangle = 0$. As $R(T) = V$, this equivalent to $\langle f, v\rangle = 0$ for all $v \in V$, i.e., $f = 0$. □

Combination of Lemma 2.1 and 2.2, we obtain a useful criteria for the operator T to be surjective.

Corollary 2.3. *Let U and V be Banach spaces and let $T \in \mathscr{B}(U,V)$. Then T is surjective if and only if T' is bounded below, i.e. $\|T'f\| \geq c\|f\|$ for all $f \in V'$.*

2.2. **Abstract Variational Problems.** Let

$$a(\cdot,\cdot) : U \times V \mapsto \mathbb{R}$$

be a bilinear form on two Banach spaces U and V, i.e., it is linear to each variable. It will introduce two linear operators

$$A : U \mapsto V', \ \text{ and } \ A' : V \mapsto U'$$
$$\text{by} \quad \langle Au, v\rangle = \langle u, A'v\rangle = a(u,v).$$

We consider the operator equation: given an $f \in V'$, find $u \in U$ such that

$$Au = f \qquad \text{in } V', \tag{6}$$

or equivalently

$$a(u, v) = \langle f, v \rangle \quad \text{for all } v \in V.$$

To begin with, we assume both A and A' are continuous which can be derived from the continuity of the bilinear form.

(C) The bilinear form $a(\cdot, \cdot)$ is continuous in the sense that

$$a(u, v) \leq C\|u\|_U \|v\|_V, \quad \text{for all } u \in U, v \in V.$$

The minimal constant satisfies the above inequality will be denoted by $\|a\|$. With this condition, it is easy to check that A and A' are bounded operators and $\|A\|_{U \to V'} = \|A'\|_{V \to U'} = \|a\|$. Later on, if we do not emphasize the norm dependence, we shall skip the subscript in $\| \cdot \|$ which should be clear from the context.

The following conditions discuss the existence and the uniqueness.

(E)

$$\inf_{v \in V} \sup_{u \in U} \frac{a(u, v)}{\|u\|\|v\|} = \alpha_E > 0.$$

(U)

$$\inf_{u \in U} \sup_{v \in V} \frac{a(u, v)}{\|u\|\|v\|} = \alpha_U > 0.$$

Theorem 2.4. *Assume the bilinear form $a(\cdot, \cdot)$ is continuous, i.e., (C) holds, the problem (6) is well-posed if and only if (E) and (U) hold. Furthermore if (E) and (U) hold, then*

$$\|A^{-1}\| = \|(A')^{-1}\| = \alpha_U^{-1} = \alpha_E^{-1} = \alpha^{-1},$$

and thus for the solution to $Au = f$

$$\|u\|_U \leq \frac{1}{\alpha} \|f\|_{V'}.$$

Proof. We can interpret (E) as $\|A'v\| \geq \alpha_E \|v\|$ for all $v \in V$ which is equivalent to A is surjective. Similarly (U) is $\|Au\| \geq \alpha_U \|u\|$ which is equivalent to A is injective. So $A : U \to V$ is bijective and by open mapping theorem, A^{-1} is bounded and it is not hard to prove the norm is α_U^{-1}. Proof for A' is similar. □

Let us take the inf-sup condition (E) as an example to show how to verify (E). It is easy to show that (E) is equivalent to

$$\text{for any } v \in V, \text{ there exists } u \in U, \text{ s.t. } a(u, v) \geq \alpha \|u\| \|v\|. \tag{7}$$

We shall present a slightly different characterization of (E). With this characterization, it is transformed to a construction of a suitable function.

Proposition 2.5. *The inf-sup condition (E) is equivalent to that for any $v \in V$, there exists $u = u(v) \in U$, such that*

$$a(u, v) \geq C_1 \|v\|^2, \quad \textit{and } \|u\| \leq C_2 \|v\|. \tag{8}$$

Proof. Obviously (8) will imply (7) with $\alpha = C_1/C_2$. We now prove (E) implies (8). For any $v \in V$, by Corollary 1.3, there exists $f \in V'$ s.t. $f(v) = \|v\|^2$ and $\|f\| = \|v\|$. Since A is onto, we can find u s.t. $Au = f$ and by open mapping theorem, we can find a u with $\|u\| \leq \alpha_E^{-1} \|f\| = \alpha_E^{-1} \|v\|$ and $a(u, v) = \langle Au, v \rangle = f(v) = \|v\|^2$. □

For a given v, the desired u satisfying (8) could dependent on v in a subtle way. A special and simple case is $u = v$ when $U = V$. The corresponding result is known as Lax-Milgram Theorem.

Lemma 2.6 (Lax-Milgram)**.** *For a bilinear form $a(\cdot,\cdot)$ on $V \times V$, if it satisfies*

(1) Continuity: $a(u,v) \le \beta\|u\|\|v\|$;
(2) Coercivity : $a(u,u) \ge \alpha\|u\|^2$,

then for any $f \in V'$, there exists a unique $u \in V$ such that

$$a(u,v) = \langle f, v\rangle,$$

and

$$\|u\| \le 1/\alpha\|f\|.$$

The simplest case is the bilinear form $a(\cdot,\cdot)$ is symmetric and positive definite on V. Then $a(\cdot,\cdot)$ defines a new inner product. Lax-Milgram theorem is simply the Riesz representation theorem.

When $a(\cdot,\cdot)$ is symmetric not necessary positive definite, we can simplify the conditions.

Corollary 2.7 (Symmetric Operator Equation)**.** *When the bilinear form $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ is symmetric, i.e. $a(u,v) = a(v,u)$. Then if*

$$\alpha\|u\|_V \le \|Au\|_{V'} \le \beta\|u\|_V, \tag{9}$$

then for any $f \in V'$, there exists a unique $u \in V$ such that

$$a(u,v) = \langle f, v\rangle,$$

and

$$\|u\| \le 1/\alpha\|f\|.$$

2.3. **Conforming Discretization of Variational Problems.** We consider conforming discretizations of the variational problem

$$a(u,v) = \langle f, v\rangle \tag{10}$$

in the finite dimensional subspaces $U_h \subset U$ and $V_h \subset V$. Find $u_h \in U_h$ such that

$$a(u_h, v_h) = \langle f, v_h\rangle, \quad \text{for all } v_h \in V_h. \tag{11}$$

The existence and uniqueness of (11) is equivalent to the following discrete inf-sup conditions:

$$(\mathbf{D}) \qquad \inf_{u\in U_h}\sup_{v\in V_h}\frac{a(u_h,v_h)}{\|u_h\|\|v_h\|} = \inf_{v\in V_h}\sup_{u\in U_h}\frac{a(u_h,v_h)}{\|u_h\|\|v_h\|} = \alpha_h > 0.$$

With appropriate choice of basis, (11) has a matrix form. To be well defined, first of all the matrix should be square. Second the matrix should be non- singular. For a squared matrix, two inf-sup conditions are merged into one. To be uniformly stable, the constant α_h should be uniformly bounded below.

An abstract error analysis can be established using inf-sup conditions. The key property for the conforming discretization is the following Galerkin orthogonality

$$a(u - u_h, v_h) = 0, \quad \text{for all } v_h \in V_h.$$

Theorem 2.8. *If the bilinear form $a(\cdot,\cdot)$ satisfies (C), (E), (U) and (D), then there exists a unique solution $u \in U$ to (10) and a unique solution $u_h \in U_h$ to (11). Furthermore*

$$\|u - u_h\| \le \frac{\|a\|}{\alpha_h}\inf_{v_h\in U_h}\|u - v_h\|.$$

Proof. With those assumptions, we know for a given $f \in V'$, the corresponding solutions u and u_h are well defined. Let us define a projection operator $P_h : U \mapsto U_h$ by $P_h u = u_h$. Note that $P_h|_{U_h}$ is identity. In operator form $P_h = A_h^{-1} Q_h A$, where $Q_h : V' \to V_h'$ is the natural inclusion of dual spaces. We prove that P_h is a bounded linear operator and $\|P_h\| \leq \|a\|/\alpha_h$ as the following:

$$\begin{aligned}
\|u_h\| &\leq \frac{1}{\alpha_h} \sup_{v_h \in V_h} \frac{a(u_h, v_h)}{\|v_h\|} \\
&= \frac{1}{\alpha_h} \sup_{v_h \in V_h} \frac{a(u, v_h)}{\|v_h\|} \\
&\leq \frac{1}{\alpha_h} \sup_{v \in V} \frac{a(u, v)}{\|v\|} \\
&\leq \frac{\|a\|}{\alpha_h} \|u\|.
\end{aligned}$$

Then for any $w_h \in U_h$, note that $P_h w_h = w_h$,

$$\|u - u_h\| = \|(I - P)(u - w_h)\| \leq \|I - P_h\| \|u - w_h\|.$$

Since $P_h^2 = P_h$, we use the identity in [5]:

$$\|I - P_h\| = \|P_h\|,$$

to get the desired result. □

3. Inf-Sup Conditions for Saddle Point System: Brezzi Theory

3.1. **Variational problem in the mixed form.** We shall consider an abstract mixed variational problem first. Let V and P be two Banach spaces. For given $(f, g) \in V' \times P'$, find $(u, p) \in V \times P$ such that:

$$a(u, v) + b(v, p) = \langle f, v \rangle, \quad \text{for all } v \in V, \tag{12}$$

$$b(u, q) = \langle g, q \rangle, \quad \text{for all } q \in P. \tag{13}$$

Let us introduce linear operators

$$A : V \mapsto V', \text{ as } \langle Au, v \rangle = a(u, v)$$

and

$$B : V \mapsto P', B' : P \mapsto V', \text{ as } \langle Bv, q \rangle = \langle v, B'q \rangle = b(v, q).$$

Written in the operator form, the problem becomes

$$Au + B'p = f, \tag{14}$$

$$B\,u = g, \tag{15}$$

or in the block matrix form

$$\begin{pmatrix} A & B' \\ B & 0 \end{pmatrix} \begin{pmatrix} u \\ p \end{pmatrix} = \begin{pmatrix} f \\ g \end{pmatrix}. \tag{16}$$

3.2. **inf-sup conditions.** We shall study the well posedness of this abstract mixed problem (16). First we assume all bilinear forms are continuous so that all operators A, B, B' are continuous.

(C) The bilinear form $a(\cdot,\cdot)$, and $b(\cdot,\cdot)$ are continuous

$$\begin{aligned} a(u,v) &\leq C\|u\|\|v\|, \quad \text{for all } u,v \in V, \\ b(v,q) &\leq C\|v\|\|q\|, \quad \text{for all } v \in V, q \in P. \end{aligned}$$

The solvable of the second equation (15) is equivalent to B is surjective or B' is injective and $R(B')$ closed which is equivalent to the following inf-sup condition
(B)

$$\inf_{q\in P}\sup_{v\in V}\frac{b(v,q)}{\|v\|\|q\|} = \beta > 0$$

With condition (B), we have $B: V/N(B) \to P$ is an isomorphism. So given $g \in P'$, we can chose $u_1 \in V/N(B)$ such that $Bu_1 = g$ and $\|u_1\|_V \leq \beta^{-1}\|g\|_{P'}$.

After we get a unique u_1, we restrict the test function v in (12) to $N(B)$. Since $\langle v, B'q\rangle = \langle Bv, q\rangle = 0$ for $v \in N(B)$, we get the following variational form: find $u_0 \in N(B)$ such that

$$a(u_0, v) = \langle f, v\rangle - a(u_1, v), \quad \text{for all } v \in N(B). \tag{17}$$

The existence and uniqueness of u_0 is then equivalent to the two inf-sup conditions for $a(u,v)$ on space $Z = N(B)$. That is, A can be singular but restricted to $N(B)$, A is well-posed.

(A)

$$\inf_{u\in Z}\sup_{v\in Z}\frac{a(u,v)}{\|u\|\|v\|} = \inf_{v\in Z}\sup_{u\in Z}\frac{a(u,v)}{\|u\|\|v\|} = \alpha > 0.$$

After we determine a unique $u = u_0 + u_1$ in this way, we solve

$$B'p = f - Au \tag{18}$$

to get p. Since u_0 is the solution to (17), the right hand side $f - Au \in N(B)^\circ$. Thus we require $B': V \mapsto N(B)^\circ$ is an isomorphism which is also equivalent to the condition (B).

Theorem 3.1. *Assume the bilinear forms $a(\cdot,\cdot), b(\cdot,\cdot)$ are continuous, i.e., (C) holds. The mixed variational problem (16) is well-posed if and only if (A) and (B) hold. When (A) and (B) hold, we have the stability result*

$$\|u\|_V + \|p\|_P \lesssim \|f\|_{V'} + \|g\|_{P'}.$$

The following characterization of the inf-sup condition for the operator B is useful. The verification is again transferred to a construction of a suitable function. The proof is similar to that in Theorem 2.5 and thus skipped here.

Theorem 3.2. *The inf-sup condition (B) is equivalent to that: for any $q \in P$, there exists $v \in V$, such that*

$$b(v,q) \geq C_1\|q\|^2, \quad \text{and } \|v\| \leq C_2\|q\|. \tag{19}$$

Note that in general a construction of desirable $v = v(q)$, especially the control of norm $\|v\|$, may not be straightforward.

3.3. **Conforming Discretization.** We consider finite element approximation to the mixed problem: Find $u_h \in V_h$ and $p_h \in P_h$ such that

$$a(u_h, v_h) + b(v_h, p_h) = \langle f, v_h \rangle, \qquad \text{for all } v_h \in V_h, \tag{20}$$

$$b(u_h, q_h) = \langle g, q_h \rangle, \qquad \text{for all } q_h \in P_h. \tag{21}$$

We shall mainly consider the conforming case $V_h \subset V$ and $P_h \subset P$. We denote $B_h : V_h \to P_h'$ which can be written as $B_h = Q_h B I_h$ with natural embedding $I_h : V_h \hookrightarrow V$ and $Q_h : P' \hookrightarrow P_h'$, and denote $Z_h = N(B_h)$. Recall that $Z = N(B)$. In the application to Stokes equations $B = -\operatorname{div}$, the null space Z is called the divergence free space and Z_h is the discrete divergence free space. Discrete divergence free may not be divergence free.

Remark 3.3. In general $Z_h \not\subset Z$. Namely a discrete divergence free function may not be exactly divergence free. Just compare the meaning of $B_h u_h = 0$ in $(P_h)'$

$$\langle B_h u_h, q_h \rangle = 0, \quad \text{for all } q_h \in P_h,$$

with $Bu_h = 0$ in P'

$$\langle B u_h, q \rangle = 0, \quad \text{for all } q \in P.$$

If we can identify $P = P'$ and $P_h = (P_h)'$ using Riesz representation theorem, then $N(B_h) \in (P_h)^{\perp}$ which may contains non-trivial elements in P. Namely it is possible that $Bu_h \in \ker(Q_h) \cap B(V_h)$. To enforce $Z_h \subset Z$, it suffices to have $B(V_h) \subset P_h$. Indeed when $B(V_h) \subset P_h$, $Q_h B u_h = B u_h$ and thus $B_h u_h = 0$ implies $B u_h = 0$. $\Box$

The discrete inf-sup conditions for the finite element approximation will be
(D)

$$(\mathbf{A_h}) \qquad \inf_{u_h \in Z_h} \sup_{v_h \in Z_h} \frac{a(u_h, v_h)}{\|u_h\|_V \|v_h\|_V} = \alpha_h > 0,$$

$$(\mathbf{B_h}) \qquad \inf_{q_h \in P_h} \sup_{v_h \in V_h} \frac{b(v_h, q_h)}{\|v_h\|_V \|q_h\|_P} = \beta_h > 0.$$

Theorem 3.4. *If (A), (B), (C) and (D) hold, then the discrete problem is well-posed and*

$$\|u - u_h\|_V + \|p - p_h\|_P \le C \inf_{v_h \in V_h, q_h \in P_h} \|u - v_h\|_V + \|p - q_h\|_P.$$

Exercise 3.5. Let $U = V \times P$ and rewrite the mixed formulation using one bilinear form defined on U. Then use Babuška theory to prove the above theorem. Write explicitly how the constant C depends on the constants in all inf-sup conditions.

3.4. **Fortin operator.** Note that the inf-sup condition (B) in the continuous level implies: for any $q_h \in P_h$, there exists $v \in V$ such that $b(v, q_h) \ge \beta \|v\|_V \|q_h\|_P$ and $\|v\| \le C \|q_h\|$. For the discrete inf-sup condition, we need a $v_h \in V_h$ satisfying such property. One approach is to use the so-called Fortin operator [2] to get such a v_h from v.

Definition 3.6 (Fortin operator)**.** *A linear operator* $\Pi_h : V \to V_h$ *is called a* Fortin operator *if*

(1) $b(\Pi_h v, q_h) = b(v, q_h)$ *for all* $q_h \in P_h$
(2) $\|\Pi_h v\|_V \le C \|v\|_V$.

Theorem 3.7. *Assume the inf-sup condition (B) holds and there exists a Fortin operator* Π_h*, then the discrete inf-sup condition* (B_h) *holds.*

Proof. The inf-sup condition (B) in the continuous level implies: for any $q_h \in P_h$, there exists $v \in V$ such that $b(v, q_h) \geq \beta \|v\| \|q_h\|$ and $\|v\| \leq C\|q_h\|$. We choose $v_h = \Pi_h v$.

By the definition of Fortin operator

$$b(v_h, q_h) = b(v, q_h) \geq \beta \|v\|_V \|q_h\|_P \geq \beta C \|v_h\|_V \|q_h\|_P.$$

The discrete inf-sup condition then follows. □

4. EXERCISE

1. Consider the case A is symmetric and positive definite (SPD). Define $S = BA^{-1}B'$ which is the Schur complement of A.
   (a) Prove that if B is surjective, S is also SPD.
   (b) Estimate the eigenvalue of $\mathcal{L} := (A, B'; B, O)$ in the inner product defined by $\mathcal{A} = (A, S)$ of $V \times P$. That is consider generalized eigenvalue problem for the saddle point system

$$Ax + B'y = \lambda Ax, \tag{22}$$

$$Bx \qquad = \lambda Sy. \tag{23}$$

   and give bound for λ in term of the constant β in the inf-sup condition (**B**).
   (c) The stability can be given by the spectral radius of $\mathcal{A}$ as $\mathcal{A}$ is symmetric. Prove that the solution (u, p) to the saddle point system $\mathcal{L}(u, p) = (f, g)$ satisfies

$$\|(u, p)\|_{\mathcal{A}} \leq \frac{2}{\sqrt{1 + 4\beta^2} - 1} \|(f, g)\|_{\mathcal{A}^{-1}}. \tag{24}$$

2. Consider $\mathcal{L} := (A, B'; B, -C)$, where A and C are semi-symmetric and positive definite. Prove that the well-posedness of the operator equation $\mathcal{L}(u, p) = (f, g)$

$$\mu \|(u, p)\|_{V \times Q} \leq \|\mathcal{L}(u, p)\|_{V'} \leq L \|(u, p)\|_{V \times Q}, \tag{25}$$

   is equivalent to: there are constants $\underline{c}_v, \bar{c}_v, \underline{c}_q, \bar{c}_q > 0$ s.t.

$$\underline{c}_v^2 \|w\|_V^2 \leq \|Aw\|_{V'}^2 + \|Bw\|_{Q'}^2 \leq \bar{c}_v^2 \|w\|_V^2 \quad \forall w \in V, \tag{26}$$

$$\underline{c}_q^2 \|r\|_Q^2 \leq \|Cr\|_{Q'}^2 + \|B'r\|_{V'}^2 \leq \bar{c}_q^2 \|r\|_Q^2 \quad \forall r \in Q. \tag{27}$$

REFERENCES

[1] I. Babuka. The finite element method with Lagrangian multipliers. *Numer. Math.*, 20:179–192, 1973. 1

[2] M. Fortin. An analysis of the convergence of mixed finite element methods. *RAIRO Anal. Numer*, 11(R3):341–353, 1977. 12

[3] G. Strang. The Fundamental Theorem of Linear Algebra. *American Mathematical Monthly*, 100(9):848–855, 1993. 4

[4] G. Strang. *Linear Algebra and Its Applications*. New York, Academic Press, fourth edi edition, 2006. 5

[5] J. Xu and L. Zikatanov. Some Observations on {Babu{š}ka} and {Brezzi} Theories. *Numer. Math.*, 94(1):195–202, Mar. 2003. 10