1 Constrained Optimization

Let B be an $M \times N$ matrix, a linear operator from the space $\mathbb{R}^N$ to $\mathbb{R}^M$ with adjoint B^T. For (column vectors) $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$ we have $\mathbf{x} \cdot B^T\mathbf{y} = B\mathbf{x} \cdot \mathbf{y}$. This vanishes for all $y \in \mathbb{R}^M$ exactly when $\mathbf{x} \in \text{Ker}(B)$, and this is equivalent to having $\mathbf{x}$ perpendicular to $\text{Rg}(B^T)$, the range of B^T. But this is the column space of B^T which is the row space of B, and $\mathbf{x}$ perpendicular to the rows of B is equivalent to $B\mathbf{x} = \mathbf{0}$. Thus we have $\text{Ker}(B) = \text{Rg}(B^T)^\perp$, the orthogonal complement of the range of B^T, and it follows that $\text{Ker}(B)^\perp = \text{Rg}(B^T)$ and $\text{Ker}(B^T)^\perp = \text{Rg}(B)$. We begin by characterizing those linear operators on more general spaces with this property.

Let V and W be Hilbert spaces, and let $\mathcal{B} : V \to W'$ be continuous and linear. Define the *adjoint* operator $\mathcal{B}' : W \to V'$ by $\mathcal{B}'w(v) = \mathcal{B}v(w)$, $\forall v \in V,\ w \in W$. Then $\mathcal{B}'$ is continuous, and its adjoint is given by $\mathcal{B}'' = \mathcal{B}$.

NOTE: We do not identify V and V' by the Riesz map, since this map is frequently equivalent to a boundary-value problem. However, We will identify V and V'', since this involves the composition of the Riesz map followed by its inverse.

Let U be a subset of W. The *annihilator* of U is the set of functionals given by

$$U^0 \equiv \{f \in W' : \ f(w) = 0 \ \forall w \in U\}.$$

Then it follows that U^0 is a closed subspace of W'.

The Inf-Sup Condition

Let the continuous and linear operator $\mathcal{B} : V \to W'$ be given as above. A direct computation shows that

$$\begin{aligned}
(\text{Ker}\,\mathcal{B}')^0 &= \{f \in W' : \ f(w) = 0 \ \forall w \in \text{Ker}\,\mathcal{B}'\} \\
&= \{f \in W' : \ f(w) = 0 \ \forall w : \mathcal{B}'w(v) = 0 \ \forall v \in V\} \\
&= \{f \in W' : \ f(w) = 0 \ \forall w : \mathcal{B}v(w) = 0 \ \forall v \in V\} \\
&\supset \mathcal{B}(V) = \text{Rg}\,\mathcal{B}.
\end{aligned}$$

so, we have $(\operatorname{Ker}\mathcal{B}')^0 \supset \overline{\operatorname{Rg}\mathcal{B}}$. Suppose $f_0 \in W'$ but $f_0 \notin \overline{\operatorname{Rg}\mathcal{B}}$. Then the *separation theorem* gives a $w^* \in W$ and $a \in \mathbb{R}$ such that $f_0(w^*) > a$ and $\mathcal{B}v(w^*) \leq a \ \forall v \in V$, hence $a \geq 0$ and $\mathcal{B}v(w^*) = 0 \ \forall v \in V$. These show that $f_0(w^*) \neq 0$ and $\mathcal{B}'(w^*) = 0$, so $f_0 \notin (\operatorname{Ker}\mathcal{B}')^0$.

Theorem 1.1. *If $\mathcal{B} : V \to W'$ is continuous and linear, then the closure of the range of $\mathcal{B}$ is the annihilator of the kernel of $\mathcal{B}'$, that is,*

$$\overline{\mathcal{B}(V)} = (\operatorname{Ker}\mathcal{B}')^0.$$

Corollary 1.2. *There exists a $\beta > 0$ such that*

$$\|\mathcal{B}(v)\|_{W'} \geq \beta\|v\|_V \ \forall v \in V, \tag{1}$$

if and only if $\mathcal{B}$ is an isomorphism of V onto $(\operatorname{Ker}\mathcal{B}')^0$.

In this case, we say that $\mathcal{B}$ is *bounding.*

Proof. From (1), we see that $\mathcal{B}$ is injective and $\mathcal{B}^{-1}$ is continuous, hence, that the range of $\mathcal{B}$ is closed. □

Remark 1.1. *The condition* (1) *is equivalent to*

$$\sup_{w\in W} \frac{\mathcal{B}v(w)}{\|w\|_W} \geq \beta\|v\|_V \ \forall v \in V,$$

and this is precisely the inf-sup condition

$$\inf_{v\in V} \sup_{w\in W} \frac{\mathcal{B}v(w)}{\|v\|_V\|w\|_W} \geq \beta > 0. \tag{2}$$

Also, it follows easily that when additionally the adjoint $\mathcal{B}'$ is injective, the operator $\mathcal{B}$ is an isomorphism onto W'.

Corollary 1.3. *The linear $\mathcal{B} : V \to W'$ is an isomorphism if and only if it satisfies*

- *$\mathcal{B}$ is bounded: there is a constant C_B such that*

$$|\mathcal{B}v(w)| \leq C_B\|v\|_V\|w\|_W, \quad v \in V,\ w \in W, \tag{3}$$

- *$\mathcal{B}$ is bounding:* (1) *holds for some $\beta > 0$, and*

- *for every $w \in W$, $w \neq 0$, there is a $v \in V$ with $\mathcal{B}v(w) \neq 0$.*

Proof. The last condition implies that $\operatorname{Ker}\mathcal{B}' = \{0\}$ so $\overline{\mathcal{B}(V)} = W'$ by Theorem 1.1, and Corollary 1.2 shows $\mathcal{B}$ is an isomorphism. The reverse implication is clear. □

Finally, we apply Corollary 1.2 to $\mathcal{B}'$ to obtain the equivalence of the first two parts of the following.

Theorem 1.4. *Assume $\mathcal{B} : V \to W'$ is continuous and linear. The following are equivalent:*

- *The adjoint $\mathcal{B}' : W \to (\operatorname{Ker}\mathcal{B})^0$ is an isomorphism.*
- *$\mathcal{B}'$ is bounding:*

$$\inf_{w \in W} \sup_{v \in V} \frac{\mathcal{B}'w(v)}{\|v\|_V \|w\|_W} \geq \beta > 0. \tag{4}$$

- *The restriction to the orthogonal complement $\mathcal{B} : (\operatorname{Ker}\mathcal{B})^\perp \to W'$ is an isomorphism, and*

$$\|\mathcal{B}v\|_{W'} \geq \beta \|v\|_V, \quad v \in (\operatorname{Ker}\mathcal{B})^\perp. \tag{5}$$

- *$\mathcal{B} : V \to W'$ is a surjection.*

Proof. If the first condition holds, then for some $\beta > 0$ we have

$$\|\mathcal{B}'w\|_{V'} \geq \beta \|w\|_W, \quad w \in W. \tag{6}$$

Thus, for each $v \in (\operatorname{Ker}\mathcal{B})^\perp$ we pick $w \in W$ with $\mathcal{B}'w(z) = (v, z)_V,\ z \in V$. From (6) we get $\|v\|_V = \|\mathcal{B}'w\|_{V'} \geq \beta\|w\|_W$. Setting $z = v$ above we obtain the equality in

$$\sup_{u \in W} \frac{\mathcal{B}v(u)}{\|u\|_W} \geq \frac{\mathcal{B}v(w)}{\|w\|_W} = \frac{(v, v)_V}{\|w\|_W} \geq \beta \|v\|_V. \tag{7}$$

This shows $\mathcal{B} : (\operatorname{Ker}\mathcal{B})^\perp \to W'$ satisfies the conditions of Corollary 1.2, so it is an isomorphism.

Assume $\mathcal{B} : (\operatorname{Ker}\mathcal{B})^\perp \to W'$ is an isomorphism and (5) holds. Then we have

$$\|w\|_W = \sup_{v \in \operatorname{Ker}\mathcal{B}^\perp} \frac{\mathcal{B}'w(v)}{\|\mathcal{B}v\|_{W'}} \leq \sup_{v \in \operatorname{Ker}\mathcal{B}^\perp} \frac{\mathcal{B}v(w)}{\beta\|v\|_V}. \tag{8}$$

This shows that (4) is satisfied, so the first three items are equivalent.

Finally, note that each $v \in V$ can be written $v = v_0 + v_\perp$ with $v_0 \in \operatorname{Ker}\mathcal{B}$ and $v_\perp \in (\operatorname{Ker}\mathcal{B})^\perp$ and $\|v_\perp\|_V = \inf\{\|v+z\|_V : z \in \operatorname{Ker}\mathcal{B}\}$. Thus, the third and fourth parts are equivalent by the open mapping theorem. $\square$

Remark 1.2. *In this case, the estimate*

$$\|\mathcal{B}v\|_{W'} \geq \beta \inf_{z\in\operatorname{Ker}\mathcal{B}} \|v+z\|_V, \quad v \in V, \tag{9}$$

holds, and (9) *can be written in terms of the norm*

$$\|\tilde{v}\|_{V/\operatorname{Ker}\mathcal{B}} = \inf_{z\in\operatorname{Ker}\mathcal{B}} \|v+z\|_V$$

on the quotient space $\tilde{V} = V/\operatorname{Ker}\mathcal{B}$.

The Closed Range Theorem

Lemma 1.5. *If $\mathcal{B} : V \to W'$ has closed range,* $\operatorname{Rg}\mathcal{B}$*, then its adjoint $\mathcal{B}' : W \to V'$ has closed range,* $\operatorname{Rg}\mathcal{B}'$*.*

Proof. Suppose $\mathcal{B}_1$ is just $\mathcal{B}$ regarded as an operator from V to the Hilbert (sub)space $\operatorname{Rg}\mathcal{B} = \overline{\operatorname{Rg}\mathcal{B}}$. To compute the adjoint $(\mathcal{B}_1)' : (\operatorname{Rg}\mathcal{B})' \to V'$, let $g_1 \in (\operatorname{Rg}\mathcal{B})'$ and note that it can be extended to a $g \in W'' = W$. Then for each $v \in V$ we have $(\mathcal{B}_1)'g_1(v) = \langle \mathcal{B}_1 v, g_1\rangle = (\mathcal{B}v, g) = \mathcal{B}'g(v)$, so $(\mathcal{B}_1)'(g_1) = \mathcal{B}'(g)$, and we have shown that $\operatorname{Rg}\mathcal{B}_1' = \operatorname{Rg}\mathcal{B}'$. Hence, it suffices to assume $\operatorname{Rg}\mathcal{B} = W'$. But if $\mathcal{B}$ is a surjection, then $\mathcal{B}'$ is bounding and has a closed range. $\square$

This applies as well to $\mathcal{B}'$, so we obtain the following summary.

Theorem 1.6. *The following are equivalent:*

- $\mathcal{B} : V \to W'$ *has closed range*
- $\mathcal{B}(V) = (\operatorname{Ker}\mathcal{B}')^0$
- $\sup_{u\in W} \frac{\mathcal{B}v(u)}{\|u\|_W} \geq \beta \inf_{z\in\operatorname{Ker}\mathcal{B}} \|v+z\|_V, \quad v \in V,$
- *the adjoint* $\mathcal{B}' : W \to V'$ *has closed range*
- $\mathcal{B}'(W) = (\operatorname{Ker}\mathcal{B})^0$

- $\sup_{v\in V} \frac{\mathcal{B}v(w)}{\|v\|_V} \geq \beta \inf_{z\in \operatorname{Ker}\mathcal{B}'} \|w+z\|_V, \quad w \in W,$

A very special but important case of Corollary 1.3 is the following.

Corollary 1.7. *Assume that the linear operator $\mathcal{A} : V \to V'$ is continuous and* V-coercive*: there is an $\alpha > 0$ such that*

$$\mathcal{A}v(v) \geq \alpha \|v\|_V^2, \quad v \in V. \tag{10}$$

Then $\mathcal{A} : V \to V'$ is an isomorphism: for each $f \in V'$ there is a unique $u \in V$ for which $\mathcal{A}u = f$, and moreover $\|u\|_V \leq \alpha^{-1}\|f\|_{V'}$.

This shows the equation $\mathcal{A}u = f$, *i.e.*,

$$u \in V : \quad \mathcal{A}u(v) = f(v) \ \forall v \in V, \tag{11}$$

is *well-posed*: for each $f \in V'$, (11) has a unique solution that depends continuously on f.

The Mixed Formulation

Let V and W be Hilbert spaces, and let $\mathcal{A} : V \to V'$ and $\mathcal{B} : V \to W'$ be continuous and linear. Let $f \in V'$ and $g \in \mathcal{B}(V) \subset W'$ be given.

Suppose that $\mathcal{A}$ is symmetric and non-negative. Then the equation (11) characterizes the solution of the *minimization problem*

$$u \in V : \quad \tfrac{1}{2}\mathcal{A}u(u) - f(u) \leq \tfrac{1}{2}\mathcal{A}v(v) - f(v) \ \forall v \in V, \tag{12}$$

If we minimize subject to a *constraint*, $\mathcal{B}v = g$, then this becomes

$$u \in V,\ \mathcal{B}u = g : \\ \tfrac{1}{2}\mathcal{A}u(u) - f(u) \leq \tfrac{1}{2}\mathcal{A}v(v) - f(v) \ \forall v \in V \text{ with } \mathcal{B}v = g. \tag{13}$$

We shall assume $g \in \operatorname{Rg}(\mathcal{B})$. The set $\{v \in V : \ \mathcal{B}v = g\}$ is convex, in fact, a translate of $\operatorname{Ker}\mathcal{B}$, so the constrained minimum is characterized by

$$u \in V : \quad \mathcal{B}u = g \text{ and } \mathcal{A}u - f \in (\operatorname{Ker}\mathcal{B})^0. \tag{14}$$

Consider the problem (14) without assuming $\mathcal{A}$ is symmetric. If $\mathcal{A}$ is V-coercive on $\operatorname{Ker}\mathcal{B}$, then there exists a unique solution. To see this, let $u_g \in V$ with $\mathcal{B}u_g = g$. Then we seek $u_0 = u - u_g$ for which

$$u_0 \in \operatorname{Ker}\mathcal{B} : \ \mathcal{A}u_0 + \mathcal{A}u_g - f \in (\operatorname{Ker}\mathcal{B})^0.$$

But Corollary 1.7 applied to the restriction of $\mathcal{A}$ to $\operatorname{Ker}\mathcal{B}$ shows there is exactly one such u_0. Thus, there is a unique solution to the constrained equation

$$u \in V : \quad \mathcal{B}u = g \text{ and } \mathcal{A}u(v) = f(v) \ \forall v \in \operatorname{Ker}\mathcal{B}. \tag{15}$$

Note that the addition of the constraint equation for $\mathcal{B}$ corresponds to a relaxation of the equation for $\mathcal{A}$.

If additionally $\mathcal{B}'$ is bounding, then $\mathcal{A}u - f$ belongs to the range of $\mathcal{B}'$, so there exists a unique solution of the *mixed formulation*, a pair

$$[u, p] \in V \times W : \ \mathcal{B}u = g \text{ and } \mathcal{A}u + \mathcal{B}'p = f \,.$$

The vector $p \in W$ realizes the relaxation of the equation in V', and we shall see that it is the *Lagrange multiplier* for the constraint, $\mathcal{B}u = g$.

Theorem 1.8. *Assume that the linear operators $\mathcal{A} : V \to V'$, $\mathcal{B} : V \to W'$ are continuous from the indicated Hilbert spaces V, W to their duals, and*

- *$\mathcal{A}$ is non-negative and it is V-coercive on* $\operatorname{Ker}\mathcal{B}$*: there is an $\alpha > 0$ such that*
$$\mathcal{A}v(v) \geq \alpha \|v\|_V^2, \quad v \in \operatorname{Ker}\mathcal{B}. \tag{16}$$
- *$\mathcal{B}'$ is bounding,* i.e., *it is injective and*
$$\inf_{q \in W} \sup_{v \in V} \frac{|\mathcal{B}v(q)|}{\|v\|_V \|q\|_W} \geq \beta > 0.$$

Then for every $f \in V'$ and $g \in W'$ the mixed system

$$\begin{aligned} u \in V, \ p \in W : \ \mathcal{A}u + \mathcal{B}'p &= f \in V', \\ \mathcal{B}u &= g \in W', \end{aligned} \tag{17}$$

has a unique solution in $V \times W$, and it satisfies the estimate

$$\|u\|_V + \|p\|_W \leq K(\|f\|_{V'} + \|g\|_{W'}). \tag{18}$$

Proof. We have shown above that the system (17) has a unique solution. Let's obtain the precise form of the estimate (18). First, u_g can be chosen from $(\operatorname{Ker}\mathcal{B}')^\perp$ with $\|u_g\|_V \leq \frac{1}{\beta}\|g\|_{W'}$. Then u_0 is obtained with the estimate $\|u_0\|_V \leq \frac{1}{\alpha}(\|f\|_{V'} + \|\mathcal{A}u_g\|_{V'})$. Finally, p satisfies $\|p\|_W \leq \frac{1}{\beta}(\|f\|_{V'} + \|\mathcal{A}u\|_{V'})$. Combining these, we find

$$\|u\|_V \leq \frac{1}{\alpha}(\|f\|_{V'} + \frac{1}{\beta}(\|\mathcal{A}\|_{\mathcal{L}(V,V')} + \alpha)\|g\|_{W'}) \tag{19}$$

$$\|p\|_W \leq \frac{1}{\beta}(\|f\|_{V'} + \|\mathcal{A}\|_{\mathcal{L}(V,V')}\|u\|_V) \tag{20}$$

□

Here's the result from [1] with an additional operator.

Theorem 1.9. *Assume that the linear operators $\mathcal{A} : V \to V'$, $\mathcal{B} : V \to W'$, $\mathcal{C} : W \to W'$ are continuous from the indicated Hilbert spaces V, W to their duals, and*

- *$\mathcal{A}$ is non-negative and V-coercive on* $\operatorname{Ker}\mathcal{B}$*: there is an $\alpha > 0$ such that*
$$\mathcal{A}v(v) \geq \alpha\|v\|_V^2, \quad v \in \operatorname{Ker}\mathcal{B}. \tag{21}$$
- *$\mathcal{C}$ is non-negative, symmetric, and*
- *$\mathcal{B}'$ is bounding,* i.e., *it is injective and*
$$\inf_{q\in W}\sup_{v\in V}\frac{|\mathcal{B}v(q)|}{\|v\|_V\|q\|_W} \geq \beta > 0.$$

Then for every $f \in V'$ and $g \in W'$ the system

$$\begin{aligned} u \in V,\ p \in W :\ \mathcal{A}u + \mathcal{B}'p &= f \in V', \\ -\mathcal{B}u + \mathcal{C}p &= g \in W', \end{aligned} \tag{22}$$

has a unique solution in $V \times W$, and it satisfies the estimate

$$\|\mathbf{u}\|_V + \|p\|_W \leq K(\|f\|_{V'} + \|g\|_{W'}). \tag{23}$$

Here's the more general case from [2].

Theorem 1.10. *Assume that the linear operators $\mathcal{A}, \mathcal{B}, \mathcal{C}$ are continuous as indicated from the Hilbert spaces V, W to their duals, and*

- $\mathcal{A}$ *is non-negative and invertible on* $\mathrm{Ker}\,\mathcal{B}$, i.e.,

$$\inf_{\mathbf{u}\in\mathrm{Ker}\,\mathcal{B}} \sup_{\mathbf{v}\in\mathrm{Ker}\,\mathcal{B}} \frac{\mathcal{A}\mathbf{u}(\mathbf{v})}{\|\mathbf{u}\|_V\|\mathbf{v}\|_V} \ge c_0 > 0,$$
$$\inf_{\mathbf{v}\in\mathrm{Ker}\,\mathcal{B}} \sup_{\mathbf{u}\in\mathrm{Ker}\,\mathcal{B}} \frac{\mathcal{A}\mathbf{u}(\mathbf{v})}{\|\mathbf{u}\|_V\|\mathbf{v}\|_V} \ge c_0\,, \tag{24}$$

- $\mathcal{C}$ *is non-negative, symmetric, W-coercive on* $\mathrm{Ker}\,\mathcal{B}'$, i.e.,

$$\mathcal{C}q(q) \ge c_0\|q\|_W^2,\ q \in \mathrm{Ker}\,\mathcal{B}', \tag{25}$$

- $\mathcal{B}$ *has closed range,* i.e.,

$$\sup_{\mathbf{v}\in V} \frac{|\mathcal{B}\mathbf{v}(q)|}{\|\mathbf{v}\|_V} \ge c_0\|q\|_{W/\,\mathrm{Ker}\,\mathcal{B}'},$$

Then for every $f \in V'$ and $\mathbf{g} \in \mathrm{Rg}\,\mathcal{B}$ the system (22) *has a solution which is unique in $V \times W/(\mathrm{Ker}\,\mathcal{B}' \cap \mathrm{Ker}\,\mathcal{C})$, and it satisfies the estimate*

$$\|\mathbf{u}\|_V + \|p\|_{W/\,\mathrm{Ker}\,\mathcal{B}'} \le K(\|f\|_{V'} + \|g\|_{W'}). \tag{26}$$

Saddle-point and Lagrangian

In this section we shall assume that $\mathcal{A} : V \to V'$, $\mathcal{B} : V \to W'$ are continuous and linear and that $\mathcal{A}$ is symmetric and non-negative: $\mathcal{A} = \mathcal{A}' \ge 0$. Also $f \in V'$ and $g \in W'$ are given. Define the *energy functional* $J : V \to \mathrm{I\!R}$ by

$$J(v) = \tfrac{1}{2}\mathcal{A}v(v) - f(v), \quad v \in V,$$

and the *Lagrangian* $\mathcal{L} : V \times W \to \mathrm{I\!R}$ as

$$\mathcal{L}(v,q) = J(v) + \mathcal{B}v(q) - g(q), \quad [v,q] \in V \times W.$$

Note that the derivatives of $\mathcal{L}(v,q)$ with respect to v and q are given by $\mathcal{L}_v(v,q) = \mathcal{A}v - f + \mathcal{B}'q$ and $\mathcal{L}_q(v,q) = \mathcal{B}v - g$, respectively, and these vanish at $[u,p]$ when the system (17) is satisfied. This relates the constrained minimization problem to an extremum over a linear space, namely, the *saddle-point problem*

$$[u,p] \in V \times W :\ \mathcal{L}(u,q) \le \mathcal{L}(u,p) \le \mathcal{L}(v,p)\ \forall v \in V,\ q \in W. \tag{27}$$

That is, a solution of the saddle-point problem is a *critical point* of the Lagrange function.

Theorem 1.11. *The pair $[u,p] \in V \times W$ is a solution of the saddle-point problem* (27) *if and only if it is a solution of the mixed system* (17).

Proof. For each $q \in W$ the functional $v \mapsto \mathcal{L}(v,q)$ is convex, so it takes a minimum at u exactly when $\mathcal{A}u + \mathcal{B}'q = f$. The second inequality of (27) is equivalent to the first equation of (17), and the first inequality of (27) is equivalent to $(\mathcal{B}u - g)(q - p) \leq 0$ for all $q \in W$, that is, $\mathcal{B}u = g$. $\square$

If $\mathcal{A}$ is $\operatorname{Ker}\mathcal{B}$-coercive and $\mathcal{B}'$ is bounding, then there is exactly one such solution pair.

We shall denote the infimum over a set by the *minimum* 'min' when it is attained by at least one element of that set. Similarly we denote the supremum by *maximum* 'max' when it is realized on that set.

Theorem 1.12. *The pair $[u,p]$ is a solution of the saddle-point problem* (27) *if and only if*

$$\min_{v\in V}\big(\sup_{q\in W}\mathcal{L}(v,q)\big) = \max_{q\in W}\big(\inf_{v\in V}\mathcal{L}(v,q)\big), \tag{28}$$

and this quantity is equal to $\mathcal{L}(u,p)$.

Proof. Denote the upper and lower extrema by

$$\overline{\varphi}(v) = \sup_{q\in W}\mathcal{L}(v,q), \quad \underline{\varphi}(q) = \inf_{v\in V}\mathcal{L}(v,q).$$

(These can take the values $+\infty$ and $-\infty$, respectively.)

Suppose (28) and that the minimum of $\overline{\varphi}$ is attained at u, the maximum of $\underline{\varphi}$ is attained at p, and we have $\overline{\varphi}(u) = \underline{\varphi}(p)$. The definitions of $\overline{\varphi}$ and $\underline{\varphi}$ show that $\underline{\varphi}(p) \leq \mathcal{L}(u,p) \leq \overline{\varphi}(u)$, so $\overline{\varphi}(u) = \mathcal{L}(u,p) = \underline{\varphi}(p)$ and $[u,p]$ is a saddle-point.

Note that since $\mathcal{L}(v,q) \leq \overline{\varphi}(v)$ for all $v \in V, q \in W$, we have $\underline{\varphi}(q) \leq \inf_{v\in V}\overline{\varphi}(v)$ and hence the inequality

$$\sup_{q\in W}\underline{\varphi}(q) \leq \inf_{v\in V}\overline{\varphi}(v) \tag{29}$$

always holds. If $[u,p]$ is a saddle-point, then $\overline{\varphi}(u) = \mathcal{L}(u,p) = \underline{\varphi}(p)$, so we have

$$\inf_{v\in V}\overline{\varphi}(v) \leq \overline{\varphi}(u) = \underline{\varphi}(p) \leq \sup_{q\in W}\underline{\varphi}(q).$$

Combining this with (29), we obtain

$$\inf_{v\in V} \overline{\varphi}(v) = \overline{\varphi}(u) = \mathcal{L}(u,p) = \underline{\varphi}(p) = \sup_{q\in W} \underline{\varphi}(q),$$

and this yields (28). □

Note that the upper extrema above is given by

$$\overline{\varphi}(v) = \sup_{q\in W} \mathcal{L}(v,q) = \begin{cases} J(v) \text{ if } \mathcal{B}v = g, \\ +\infty \text{ if } \mathcal{B}v \neq g, \end{cases}$$

so it follows that

$$\inf_{v\in V} \big(\sup_{q\in W} \mathcal{L}(v,q) \big) = \inf_{\mathcal{B}v=g} J(v).$$

If $[u,p]$ is a saddle-point, the inf-sup equality

$$\mathcal{L}(u,p) = \min_{v\in V} \big(\sup_{q\in W} \mathcal{L}(v,q) \big)$$

shows that

$$J(u) = \min_{\mathcal{B}v=g} J(v),$$

that is, the first component u of a saddle point is characterized as a solution of the constrained minimization problem (13). This is the *primal problem* where we began.

Let's consider the sup-inf equation

$$\mathcal{L}(u,p) = \underline{\varphi}(p) = \max_{q\in W} \big(\inf_{v\in V} \mathcal{L}(v,q) \big). \tag{30}$$

In order to characterize this equality, we assume in addition that $\mathcal{A}$ is V-coercive. Then for each $q \in W$ there is a unique solution v_q of

$$v_q \in V : \ \mathcal{A}v_q + \mathcal{B}'q = f,$$

that is, v_q is the solution of the minimization problem

$$\underline{\varphi}(q) = \mathcal{L}(v_q, q) = \inf_{v\in V} \mathcal{L}(v,q), \ q \in W.$$

Since $v_p = u$, we have $\mathcal{L}(v_p,p) = \max_{q\in W} \mathcal{L}(v_q,q)$. The definitions of $\mathcal{L}$ and v_q show that

$$\mathcal{L}(v_q,q) = \tfrac{1}{2}\mathcal{A}v_q(v_q) - f(v_q) + \mathcal{B}v_q(q) - g(q) = -\tfrac{1}{2}\mathcal{A}v_q(v_q) - g(q),$$

so we see that the function defined by

$$K(q) \equiv \tfrac{1}{2}\mathcal{A}v_q(v_q) + g(q), \quad q \in W,$$

is convex (since $q \mapsto v_q$ is affine) and it is minimized at p, that is,

$$p \in W : \; K(p) = \min_{q\in W} K(q). \tag{31}$$

This is the *dual problem.*

In order to characterize a solution of the dual problem (31), we compute the derivative $K'(p)$ from the expansion

$$\begin{aligned} K(q) &= \tfrac{1}{2}(f - \mathcal{B}'q)\mathcal{A}^{-1}((f - \mathcal{B}'q) + g(q) \\ &= \tfrac{1}{2}f\mathcal{A}^{-1}(f) - f\mathcal{A}^{-1}\mathcal{B}'q + \tfrac{1}{2}\mathcal{B}'q\mathcal{A}^{-1}\mathcal{B}'q + g(q) \end{aligned}$$

and then use the definition of v_q to obtain in turn

$$\begin{aligned} K'(p)(q) &= (g - f\mathcal{A}^{-1}\mathcal{B}')(q) + \mathcal{B}'p\mathcal{A}^{-1}\mathcal{B}'q \\ &= -\mathcal{A}u(p)\mathcal{A}^{-1}\mathcal{B}'q + g(q) = -\mathcal{B}'q(u(p)) + g(q). \end{aligned}$$

Thus we have

$$K'(p) = -\mathcal{B}u(p) + g = -\mathcal{B}\mathcal{A}^{-1}(f - \mathcal{B}'p) + g,$$

and the solution p of the dual problem (31) is characterized by the equation

$$p \in W : \; \mathcal{B}\mathcal{A}^{-1}(\mathcal{B}'p - f) = -g \text{ in } W'. \tag{32}$$

(Of course, we could obtain this directly from (17) since $\mathcal{A}$ is invertible.)

In summary, $[u, p]$ is a solution of the saddle-point problem (27), and this is equivalent to the mixed system (17), u is a solution of the primal problem (13) with constraint, and p is a solution of the dual problem (31). Also, u and p can be obtained from each other by means of the first equation of the mixed system (17) when $\mathcal{A}$ is V-coercive and $\mathcal{B}'$ is bounding.

References

[1] Dietrich Braess. *Finite elements.* Cambridge University Press, Cambridge, 1997. Theory, fast solvers, and applications in solid mechanics, Translated from the 1992 German original by Larry L. Schumaker.

[2] Franco Brezzi and Michel Fortin. *Mixed and hybrid finite element methods*, volume 15 of *Springer Series in Computational Mathematics*. Springer-Verlag, New York, 1991.