Variational Method in Hilbert Space

0.1. The Minimization Principle. Hereafter we let V denote a Hilbert space with norm $\|\cdot\|$, scalar product $(\cdot,\cdot)$, and dual space V'. A subset K of V is called *closed* if each $x_n \in K$ and $\lim x_n = x$ imply $x \in K$. The subset K is *convex* if $x, y \in K$ and $0 \le t \le 1$ imply $tx + (1-t)y \in K$. The following *minimization principle* is fundamental.

Theorem 1. *Let K be a closed, convex, non-empty subset of the Hilbert space V, and let $f \in V'$. Define $\phi(x) \equiv (1/2)\|x\|^2 - f(x)$, $x \in V$. Then there exists a unique*

$$x \in K : \phi(x) \le \phi(y) \ , \qquad y \in K \ . \tag{1}$$

Proof. Set $d \equiv \inf\{\phi(y) : y \in K\}$ and choose $x_n \in K$ such that $\lim_{n\to\infty} \phi(x_n) = d$. Then we obtain successively

$$\begin{aligned} d \le \phi\big(1/2(x_m + x_n)\big) &= (1/2)\big(\phi(x_m) + \phi(x_n)\big) - (1/8)\|x_n - x_m\|^2 \ , \\ (1/4)\|x_n - x_m\|^2 &\le \phi(x_m) + \phi(x_n) - 2d \ , \end{aligned}$$

and this last expression converges to zero. Thus $\{x_n\}$ is Cauchy, it converges to some $x \in V$ by completeness, and $x \in K$ since it is closed. Since ϕ is continuous, $\phi(x) = d$ and x is a solution of (1). If x_1 and x_2 are both solutions of (1), the last inequality shows $(1/4)\|x_1 - x_2\| \le d + d - 2d = 0$, so $x_1 = x_2$. □

The solution of the minimization problem (1) can be characterized by a *variational inequality*. For $x, y \in V$ and $t > 0$ we have

$$(1/t)\Big(\phi\big(x + t(y-x)\big) - \phi(x)\Big) = (x, y-x) - f(y-x) + (1/2)t\|y-x\|^2,$$

so the *derivative* of ϕ at x in the direction $y - x$ is given by

$$\begin{aligned} \phi'(x)(y-x) &= \lim_{t\to 0}(1/t)\Big(\phi\big(x + t(y-x)\big) - \phi(x)\Big) \\ &= (x, y-x) - f(y-x) \ . \end{aligned} \tag{2}$$

From the definition of $\phi(\cdot)$, we find that the above equals $\phi(y) - \phi(x) + (x,y) - (1/2)\|x\|^2 - (1/2)\|y\|^2$, and the CBS inequality gives

$$\phi'(x)(y-x) \le \phi(y) - \phi(x) \ , \qquad x, y \in V \ . \tag{3}$$

Suppose x is a solution of (1). Since for each $y \in K$ we have $x + t(y-x) \in K$ for small $t > 0$, it follows from (2) that

$$x \in K : \varphi'(x)(y-x) \ge 0 \ , \qquad y \in K \ .$$

Conversely, for any such x it follows from (3) that it satisfies (1). Thus, we have shown that (1) is equivalent to

$$x \in K : (x, y - x) \geq f(y - x) , \qquad y \in K . \tag{4}$$

The equivalence of (1) and (4) is merely the fact that the point where a quadratic function takes its minimum is characterized by having a non-negative derivative in each direction into the set.

0.2. Consequences of the Principle. As an example, let $x_0 \in V$ and define $f \in V'$ by $f(y) = (x_0, y)$ for $y \in V$. Then $\phi(x) = (1/2)(\|x - x_0\|^2 - \|x_0\|^2)$ so (1) means that x is that point of K which is closest to x_0. Recalling that the angle θ between $x - x_0$ and $y - x$ is determined by

$$(x - x_0, y - x) = \cos(\theta)\|x - x_0\| \; \|y - x\| ,$$

we see (4) means x is that point of K for which $-\pi/2 \leq \theta \leq \pi/2$ for every $y \in K$. We define x to be the *projection* of x_0 on K and denote it by $P_K(x_0)$.

Corollary 1. *For each closed convex non-empty* subset K *of* V *there is a projection operator* $P_K : V \to K$ *for which* $P_K(x_0)$ *is that point of* K *closest to* $x_0 \in V$*; it is characterized by*

$$P_K(x_0) \in K : \big(P_K(x_0) - x_0 \; , y - P_K(x_0)\big) \geq 0 , \qquad y \in K .$$

It follows from this characterization that the function P_K satisfies

$$\|P_K(x_0) - P_K(y_0)\|^2 \leq (P_K(x_0) - P_K(y_0), x_0 - y_0) , \qquad x_0, y_0 \in V .$$

From this we see that P_K is a *contraction*, i.e.,

$$\|P_K(x_0) - P_K(y_0)\| \leq \|x_0 - y_0\| , \qquad x_0, y_0 \in V ,$$

and that P_K satisfies the *angle condition*

$$(P_K(x_0) - P_K(y_0), x_0 - y_0) \geq 0 , \qquad x_0, y_0 \in V .$$

Corollary 2. *For each closed* subspace K *of* V *and each* $x_0 \in V$ *there is a unique*

$$x \in K : (x - x_0, y) = 0 , \qquad y \in K .$$

Two vectors $x, y \in V$ are called *orthogonal* if $(x, y) = 0$, and the *orthogonal complement* of the set S is $S^\perp \equiv \{x \in V : (x, y) = 0 \text{ for } y \in S\}$. Corollary 2 says each $x_0 \in V$ can be uniquely written in the form $x_0 = x_1 + x_2$ with $x_1 \in K$ and $x_2 \in K^\perp$ whenever K is a closed subspace. We denote this orthogonal decomposition by $V = K \oplus K^\perp$.

Exercise 1. *Show that* $S^{\perp\perp} = \overline{S}$*, the* closure *of* S*.*

The *Riesz map* $\mathcal{R}$ of V into V' is defined by $\mathcal{R}(x) = f$ if $f(y) = (x, y)$ for $y \in V$. It is easy to check that $\|\mathcal{R}x\|_{V'} = \|x\|_V$; Theorem 1 with $K = V$ shows the following by way of (4).

Corollary 3. *For each linear functional* $f \in V'$ *there is a unique vector*

$$x \in V : (x, y) = f(y) , \qquad y \in V . \tag{5}$$

Thus, the linear map $\mathcal{R}$ is *onto* V', so $\mathcal{R}$ is an isometric isomorphism of the Hilbert space V onto its dual V'.

We recognize (5) as the weak formulation of certain boundary value problems. Specifically, when $V = H_0^1$ or H^1, (5) is the Dirichlet problem (**??**) or the Neumann problem (**??**), respectively, with $c = 1$. An easy but useful generalization is obtained as follows. Let $a(\cdot,\cdot) : V \times V \to \mathbb{R}$ be bilinear (linear in each variable separately), continuous (see (**??**)), symmetric ($a(x,y) = a(y,x), x, y \in V$) and V-*elliptic*: there is a $c_0 > 0$ such that

$$a(x,x) \geq c_0 \|x\|_V^2 \ , \qquad x \in V \ . \tag{6}$$

Thus, $a(\cdot,\cdot)$ determines an *equivalent scalar product* on V: a sequence converges in V with $\|\cdot\|_V$ if and only if it converges with $a(\cdot,\cdot)^{1/2}$. Thus we may replace $(\cdot,\cdot)_V$ by $a(\cdot,\cdot)$ above.

Theorem 2. *Let $a(\cdot,\cdot)$ be a bilinear, symmetric, continuous and V-elliptic form on the Hilbert space V, let K be a closed, convex and non-empty subset of V, and let $f \in V'$. Set $\phi(x) = (\frac{1}{2})a(x,x) - f(x)$, $x \in V$. Then there is a unique*

$$x \in K : \phi(x) \leq \phi(y) \ , \qquad y \in K \ . \tag{7}$$

The solution of (7) *is characterized by*

$$x \in K : a(x, y - x) \geq f(y - x) \ , \qquad y \in K \ . \tag{8}$$

If, in addition, K is a subspace of V, then (8) *is equivalent to*

$$x \in K : a(x,y) = f(y) \ , \qquad y \in K \ . \tag{9}$$

Now (9) is precisely our weak formulation, and we see it is the special case of a *variational inequality* (8) which is the characterization of the solution of the minimization problem (7).