A GENERALIZATION OF LAPLACE’S METHOD

CHII-RUEY HWANG

ABSTRACT. Let $Q$ be Gaussian with mean 0 and covariance $B$ in a separable Hilbert space. Analogous to Laplace’s method, the weak limit (as $\theta \downarrow 0$) of $\{P_\theta \mid \theta > 0\}$, with \( \frac{dP_\theta}{dQ}(x) = C_\theta \exp\left(\frac{-H(x)}{\theta}\right) \), is considered, where
\[
H(x) = \frac{1}{2} \langle Fx, x \rangle - 2 \langle Fm, x \rangle,
\]
$F$ is s.a. nonnegative definite and bounded. If $m \in \mathfrak{H}(B^{1/2})$, then $P$ is Gaussian with mean $m - B^{1/2} \pi B^{-1/2} m$ and covariance $B^{1/2} \pi B^{1/2}$, where $\pi$ is the projection onto $\mathfrak{H}(B^{1/2}FB^{1/2})$. Moreover $P$ is the fiber measure of $Q$ on $m + \mathfrak{H}(F)$. Under stronger conditions, $P$ is induced by an affine transformation.

1. Introduction. First let us formulate Laplace’s method in a very general form and describe some known results.

Using the idea of weak convergence of probability measures, Laplace’s method can be interpreted as the following limiting procedure: as $\theta \downarrow 0$,
\[
\frac{dP_\theta}{dQ}(x) = C_\theta \exp\left(\frac{-H(x)}{\theta}\right),
\]
where $\{P_\theta \mid \theta > 0\}$ and $Q$ are probability measures on the Borel $\sigma$-algebra of a Polish space $\mathfrak{X}$, $C_\theta$ is the normalization factor and $H$ is real-valued and continuous. As in statistical mechanics, one may regard $H$ as the energy function, $\theta$ the temperature and $Q$ a fixed measure in the state space $\mathfrak{X}$ (Khinchin [7]).

The questions are: When do we have a weak limit $P$ of $P_\theta$? What is the explicit expression of $P$? Is there any intuitive interpretation of $P$ or this limiting procedure?

Let us mention some results from Hwang [6]. Let $N = \{x \mid H(x) = \inf_y H(y)\}$ denote the set of all minimal energy states. Under the assumption
\[
Q\{H(x) < a\} > 0 \quad \text{for} \quad a > \inf H(x),
\]
a necessary condition for the tightness of $\{P_\theta\}$ is $N \neq \emptyset$. If $P$ exists, it concentrates on $N$. For $Q(N) > 0$, $P$ is uniformly distributed over $N$ w.r.t. $Q$. When $Q(N) = 0$, the condition
\[
\exists \varepsilon > 0 \text{ s.t. } \{H(x) < \min H(x) + \varepsilon\} \text{ is compact}
\]
LAPLACE'S METHOD

is sufficient for the tightness of \( \{P_\theta\} \). With assumptions (A1), (B1), \( \mathcal{H} = \mathbb{R}^n \) and some smoothness conditions on \( H \) and \( Q \), \( P \) can be written in terms of the intrinsic measure on \( N \).

When \( H \) is a quadratic form, (B1) is not necessarily satisfied. Especially in the infinite dimensional case, (B1) is not quite reasonable. In this paper, we assume that \( \mathcal{H} \) is a Hilbert space with inner product \( \langle , \rangle \), \( Q \) is Gaussian with mean 0 and covariance operator \( B \); and \( H(x) = \frac{1}{2}(\langle Fx, x \rangle - 2\langle k, x \rangle) \) where \( F \) is an s.a. (selfadjoint) bounded linear operator and \( k \in \mathcal{H} \). The motivation will be explained later in this section.

Since the support of \( Q \) is \( \mathcal{R}(B) \) (the closure of the range of \( B \)) (Rajput [9]), w.l.o.g. \( B \) is assumed to be one-to-one. Hence (A1) holds. We also assume \( N \neq \emptyset \). Equivalently, \( F \) is n.d. (nonnegative definite) and \( k = F(m) \) for some \( m \).

Theorem 1 gives a sufficient condition

\[
(m + \mathcal{R}(F)) \cap \mathcal{R}(B^{1/2}) \neq \emptyset, \tag{A2}
\]

for the existence of \( P \), where \( \mathcal{R}(F) \) denotes the null space of \( F \). Also, \( P \) is Gaussian with mean and covariance provided by Theorem 1.

Before giving a geometric interpretation of \( P \), let us state some results from Krée and Tortrat [8]. Let \( X \) and \( Y \) be two closed subspaces of \( \mathcal{H} \) with \( X = Y^\perp \). Then \( Q \) can be disintegrated into \( Q(dx dy) = Q_x(dx)Q_y(dy) \), where \( Q_x \) is a Gaussian measure on \( X \) with mean 0 and covariance \( \pi_x B \pi_x \) and \( Q_y \)'s (means and covariances can be found in [8] too) are measures, which are translates of a fixed Gaussian measure on \( Y \), on the affine subspaces \( (x + Y) \). \( Q \) is defined to be the fiber measure of \( Q \) on the affine subspace \( x + Y \). In the present case let \( X = \mathcal{R}(F) \) and \( Y = \mathcal{R}(F) \). Theorem 2 gives a geometric interpretation that \( P \) is the same as \( Q_{m + \mathcal{R}(F)} \), which is the fiber measure of \( Q \) on \( \mathcal{R}(F) \) and \( Y = \mathcal{R}(F) \). Moreover, \( P \) depends on \( m + \mathcal{R}(F) \) but not on the exact form of \( F \).

Proposition 1 suggests that the limiting procedure is the same as applying an affine transform to \( Q \). Proposition 2 shows that \( (A2) \) is not a necessary condition. Finally an example is provided to show that \( (A2) \) is not redundant either.

The motivation comes from pattern theory (Grenander [3, Chapter 5], [4]). Regarding \( \exp(-H(x)/\theta) \) as the “acceptance function”, one introduces a probability measure \( P \) (frozen pattern) on \( N \) (the configuration space) via a limiting procedure. (Note that \( P \) does not depend on the exact form of \( F \).) To make things clearer, let us observe the following example of random splines without using terminologies from pattern theory. Consider the spline with knots at the integers, defined by \( Lg(t) = 0, t \notin \mathbb{Z} \), where \( L \) is a differential operator of order \( p \) and with constant coefficients. At the integers we demand that \( g \) and its first \( p - 2 \) derivatives are continuous. The curve \( r_k(t) \) which is a solution of \( Lg(t) = 0 \) on \( [k, k + 1] \) is uniquely determined by its initial condition \( G(k) = \text{column}(r_k(k), \ldots, r_k^{(p-1)}(k)) \). Let \( G(k) \) be i.i.d. Gaussian with mean zero and covariance matrix \( R \). Let

\[
V_k = \text{column}(r_k(k + 1), \ldots, r_k^{(p-2)}(k + 1))
\]
and 

\[ U_{k+1} = \text{column}(r_{k+1}(k+1), \ldots, r_{k+1}^{(p-2)}(k+1)). \]

Let us piece these (random) curves, say \( n \) pieces, together to form a spline. Then, at the integer points \( k = 1, 2, \ldots, n - 1 \), we have to “condition” on \( V_k = U_{k+1} \) in some sense (Grenander [4]). It is easy to see that there exist \( (p-1) \times p \) matrices \( B_1 \) and \( B_2 \) such that \( V_k = B_1 G(k), U_{k+1} = B_2 G(k+1) \). To calculate the “conditioned” joint distribution of \( G(1), \ldots, G(n) \), (1.1) is a reasonable choice. The density of \( P_\theta \) w.r.t. Lebesgue measure is proportional to

\[
\prod_{i=1}^{n} \exp\left(-\frac{1}{2}(a_i'R^{-1}a_i)\right) \exp\left(-\frac{1}{2\theta} \left( \sum_{i=1}^{n-1} ||B_1 a_{i+1} - B_2 a_i||^2 \right) \right),
\]

where \( a_i \) is a \( p \)-dimensional column vector. Clearly the energy function is of quadratic form (Grenander [4]).

2. Main results. Without loss of generality, we consider the infinite dimensional case only. First, we shall prove that \( P_\theta \) is Gaussian.

**Lemma 1.** The characteristic function \( \psi_\theta \) of \( P_\theta \) is

\[
\psi_\theta(t) = \exp(\langle G_\theta(F/\theta)m, t \rangle - \frac{1}{2} \langle G_\theta, t \rangle)
\]

where \( G_\theta = B^{1/2}(I + B^{1/2}(F/\theta)B^{1/2})^{-1}B^{1/2}. \)

**Proof.** It suffices to show the case \( \theta = 1 \).

Since \( B \) is one-to-one, the eigenvectors \( \{e_n\} \) of \( B \) form a c.o.n.s. (complete orthonormal set). Let \( V_n = \text{span}\{e_1, \ldots, e_n\}, \pi_n = \text{projection onto } V_n \) and \( F_n = \pi_n F \pi_n \). Define

\[
(dQ_n/dQ)(x) = C_n \exp\left( -\frac{1}{2}(\langle F_n x, x \rangle - 2\langle F_n m, x \rangle) \right);
\]

then \( Q_n \to P_1 \) weakly. Hence, the characteristic function \( \phi_n \) of \( Q_n \) converges to \( \psi_1 \).

In fact, for \( t \in V_n \)

\[
\phi_n(t) = \exp\left( i\langle (F_n + B^{-1})^{-1} F_n m, t \rangle - \frac{1}{2} \langle (F_n + B^{-1})^{-1} t, t \rangle \right)
\]

Rewrite \( (F_n + B^{-1})^{-1} \) as \( B^{1/2}(I + B^{1/2}F_nB^{1/2})^{-1}B^{1/2} \). Clearly \( (I + B^{1/2}F_nB^{1/2})^{-1} \) is bounded in \( \mathcal{C} \). By using the facts that \( \| (I + B^{1/2}F_nB^{1/2})^{-1} \| < 1, F_n \to F \) strongly and \( \cap_n V_n \) is dense,

\[
\psi_1(t) = \exp\left( i\langle B^{1/2}(I + B^{1/2}FB^{1/2})^{-1}B^{1/2}F m, t \rangle \right.
\]

\[
- \frac{1}{2} \langle B^{1/2}(I + B^{1/2}FB^{1/2})^{-1}B^{1/2} t, t \rangle \right). \quad \square
\]

The following lemma is essential for the rest of this article.

**Lemma 2.** If \( D \) is bounded n.d. and s.a., then \( (I + D/\theta)^{-1} \to \pi_D \) strongly as \( \theta \downarrow 0 \), where \( \pi_D \) is the projection onto \( \mathcal{U}(D) \).

**Proof.** Let \( E \) be the resolution of the identity for \( D \). Since the spectrum \( \sigma(D) \) is a compact subset of \([0, \infty)\), the functions \( \{\theta/(\theta + \lambda)|\theta > 0, \lambda > 0\} \) are uniformly
LAPLACE'S METHOD

bounded by 1,

\[
\frac{\theta}{\theta + \lambda} \rightarrow \begin{cases} 
0 & \text{if } \lambda \neq 0, \\
1 & \text{if } \lambda = 0, 
\end{cases}
\]

as \( \theta \downarrow 0 \).

Hence,

\[
(I + \frac{1}{\theta} D)^{-1} = \int_{\sigma(D)} \frac{\theta}{\theta + \lambda} E(d\lambda) \rightarrow \Delta E(0) = \pi_D
\]

strongly (Dunford and Schwartz [1, p. 898]). □

By Lemma 2, \( G_\theta \rightarrow B^{1/2}B^{1/2} \) strongly, where \( \pi \) is the projection onto \( \mathcal{H}(B^{1/2}FB^{1/2}) \). Obviously \( G_\theta \) and \( B^{1/2}B^{1/2} \) are bounded by the \( S \)-operator \( B \). If we can prove that \( G_\theta(F/\theta)m \) converges to some \( \hat{m} \) strongly, then \( P_\theta \rightarrow P \) weakly and \( P \) is Gaussian with mean \( \hat{m} \) and covariance operator \( B^{1/2}B^{1/2} \) (Grenander [2, p. 142]). But the convergence of \( G_\theta(F/\theta)m \) is not always true; we shall see an example later. Let us assume (A2) holds. Choose \( m \in \mathcal{H}(B^{1/2}) \) and \( m_0 \) with \( B^{1/2}(m_0) = m \). Then

\[
G_\theta(F/\theta)m = B^{1/2}m_0 - B^{1/2}(I + B^{1/2}(F/\theta)B^{1/2})^{-1}m_0 \rightarrow m - B^{1/2} \pi m_0.
\]

(2.1)

Therefore, we have

**Theorem 1.** If (A2) holds, then \( P_\theta \rightarrow P \) weakly and \( P \) is Gaussian with mean \( m - B^{1/2} \pi m_0 \) and covariance \( B^{1/2} \pi B^{1/2} \).

To relate Theorem 1 to the result in Krée and Tortrat [8], let us prove the following theorem.

**Theorem 2.** The weak limit \( P \) in Theorem 1 is the fiber measure of \( Q \) on \( m + \mathcal{H}(F) \).

**Proof.** Let \( Y = \mathcal{H}(F), X = Y^\perp = \overline{\mathcal{H}(F)}, \pi_X \) and \( \pi_Y \) denote the projections onto \( X \) and \( Y \) respectively. We know that fiber measures are translates of a fixed Gaussian measure on \( Y \) with covariance operator \( \pi_Y B^{1/2} \pi_Y = \pi_Y \pi_Y \pi_X \pi_X \pi_Y \pi_Y \); see formulae (6) and (7) in Lemma 2 of Krée and Tortrat [8]. First let us prove

\[
\pi_Y B^{1/2} \pi_Y = \pi_Y \pi_Y \pi_Y \pi_Y \pi_Y \pi_Y.
\]

Rewrite the R.H.S. as \( \pi_Y B^{1/2}(I - B^{1/2} \pi_X \pi_Y B^{1/2})^{-1} \pi_Y B^{1/2} \pi_Y \). For \( z \) with \( B^{1/2}FB^{1/2}(z) = 0 \), we have \( B^{1/2}(z) \in \mathcal{H}(F) \). Then, \( \pi_Y B^{1/2}z = 0 \) and \( (I - B^{1/2} \pi_X \pi_Y B^{1/2})^{-1} \pi_Y B^{1/2}(z) = z \).

For \( z = B^{1/2}FB^{1/2}u, \pi_X B^{1/2}(B^{1/2}FB^{1/2})u = \pi_X BFB^{1/2}u = (\pi_X \pi_Y)FB^{1/2}u \).

Hence,

\[
(I - B^{1/2} \pi_X \pi_Y B^{1/2})z = z - B^{1/2} \pi_X B^{1/2}u = z - B^{1/2}FB^{1/2}(u) = 0.
\]

Since \( \pi \) is bounded, (2.2) holds.
Now we have to relate mean $m - B^{1/2}m_0$ of $P$ to a translation $\pi_x m$ of a fixed Gaussian measure with covariance $\pi_y B^{1/2}\pi B^{1/2} \pi_y$ on $Y$. From formula (4) in Lemma 1 of Krée and Tortrat [8], we have to establish for $y \in Y$

$$
\langle y, m - B^{1/2}m_0 \rangle = \langle (\pi_x B\pi_x)^{-1}\pi_x B\pi y, \pi_x m \rangle,
$$

$$
m - B^{1/2}m_0 = m - B^{1/2}(I - B^{1/2}\pi_x (\pi_x B\pi_x)^{-1}\pi_x B^{1/2})B^{1/2}m
= B\pi_x (\pi_x B\pi_x)^{-1}\pi_x m,
$$

$$
\langle y, m - B^{1/2}m_0 \rangle = \langle y, B\pi_x (\pi_x B\pi_x)^{-1}\pi_x m \rangle
= \langle y, \pi_y B\pi_x (\pi_x B\pi_x)^{-1}\pi_x m \rangle
= \langle (\pi_x B\pi_x)^{-1}\pi_x B\pi y, \pi_x m \rangle.
$$

Hence, $P$ can be regarded as a translation $\pi_x m$ of a fixed Gaussian measure with covariance $\pi_y B^{1/2}\pi B^{1/2} \pi_y$. (Note that $\pi_x m \in X$ and $B^{1/2}\pi B^{1/2} \pi \subseteq Y$).

Let $T$ be a bounded linear operator from $\mathcal{H}$ to $\mathcal{K}$ and $m$ be a fixed element in $\mathcal{K}$. The Gaussian measure with mean $\hat{m}$ and covariance $T BT^*$ is called the induced measure of $Q$ by $m + T$. Now we consider the possibility of inducing $P$ by some $\hat{m} + T$. The obvious candidate is $B^{1/2}\pi B^{-1/2}$. By the closed graph theorem, it is not hard to show

**Proposition 1.** Under the assumptions (A2) and

$$(\mathcal{R}(F) \subseteq \mathcal{R}(B), \quad (A3)$$

$B^{-1/2}\pi B^{1/2}$ is bounded and $P$ is induced by $(m - Tm) + T$ where $T = (B^{-1/2}\pi B^{1/2})^*$ ($^*$ stands for adjoint).

For particular $F$ without assumptions (A2) and (A3), it is still possible to get similar results as in Theorem 1 and Proposition 1.

**Proposition 2.** If $F$ is of diagonal form w.r.t. $\{e_n\}$, then $P_\theta \to P$ and $P$ is Gaussian with mean $m - Tm$ and covariance operator $B^{1/2}\pi B^{1/2}$, where $T$ is the continuous extension of $B^{1/2}\pi B^{-1/2}$. Moreover $P$ is induced by $(m - Tm) + T$.

**Proof.** A direct calculation and the fact (Riesz [10, p. 301])

$$(B^{1/2}\pi B^{-1/2})^* \subseteq B^{-1/2}(B^{1/2}\pi)^* = B^{-1/2}\pi B^{1/2} \quad (2.3)$$

lead to the conclusion. □

The detailed proofs of the above two propositions can be found in Hwang [5].

**Example.** There exists an $m$ such that (2.1) does not converge and $P$ does not exist.

Let $B(e_n) = e_n/n^2$, $x = \sum_1^\infty e_n/n$ and $y = (\sum_1^\infty 1/n^2)e_1 + \sum_1^\infty (-1/n)e_n$. Then $x \bot y$. Let $z_1 = |B^{1/2}(x)|^{-1}B^{1/2}(x)$ and $\{z_n\}$ be a c.o.n.s. Clearly $F$ is well defined by $F(z_1) = 0$ and $F(z_n) = z_n$ for $n \neq 1$. $F$ is n.d. and s.a. and $\mathcal{R}(F) = \text{span}\{B^{1/2}(x)\}, y \bot \mathcal{R}(FB^{1/2})$, $x + y \in \mathcal{R}(B^{1/2})$. Let $z = x + y$; then $B^{1/2}(z) = z \pi B^{1/2}(z) = x \notin \mathcal{R}(B^{1/2})$. Thus $B^{-1/2}\pi B^{1/2}$ is not bounded.

Suppose that $G_\theta F/\theta$ converges strongly; then $B^{1/2}\pi B^{-1/2}$ is continuous. Using (2.3), we shall get that $B^{-1/2}\pi B^{1/2}$ is bounded, which is a contradiction.
Therefore, there exists $m$ such that $\{G_\theta(F/\theta)m\}$ is unbounded. Otherwise, by the principle of uniform boundedness, there will be a contradiction.

Finally, consider $P_\theta$ which is Gaussian with mean 0 and covariance $G_\theta$; then $P_\theta \to \overline{P}$, where $\overline{P}$ is Gaussian with mean 0 and covariance $B^{1/2} \pi B^{1/2}$. For any $\epsilon$ in $(0, \frac{1}{2})$, there exists a ball $B(\epsilon)$ around 0 such that $\overline{P}_\theta(B(\epsilon)) > \epsilon$ for $\theta$ sufficiently small. Therefore,

$$P_\theta(G_\theta(F/\theta)m + B(\epsilon)) > \epsilon$$

and $\{P_\theta\}$ is not tight. □

**Acknowledgement.** Most of the results in this paper came from Chapter 3 of my Ph. D. dissertation written at Brown University under the supervision of Professor Ulf Grenander. I would like to express my gratitude for his guidance. Also, I am grateful to Professor R. M. Dudley, Professor T. G. Kurtz and the referee for may helpful comments.

**References**


Institute of Mathematics, Academia Sinica, Nankang, Taipei, Taiwan, Republic of China