PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 82, Number 3, July 1981

A GENERALIZATION OF LAPLACE'S METHOD[1]

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 $\theta\downarrow 0$) P of $\{P_\theta | \theta > 0\}$, with $(dP_\theta/dQ)(x) = C_\theta \exp(-H(x)/\theta)$, is considered, where

$$H(x) = \tfrac{1}{2}(\langle Fx, x\rangle - 2\langle Fm, x\rangle),$$

F is s.a. nonnegative definite and bounded. If $m \in \mathcal{R}(B^{1/2})$, 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 π is the projection onto $\mathfrak{N}(B^{1/2}FB^{1/2})$. Moreover P is the fiber measure of Q on $m + \mathfrak{N}(F)$. 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 $\theta\downarrow 0$,

$$\frac{dP_\theta}{dQ}(x) = C_\theta \exp\left(-\frac{H(x)}{\theta}\right), \tag{1.1}$$

where $\{P_\theta | \theta > 0\}$ and Q are probability measures on the Borel σ-algebra of a Polish space $\mathcal{H}$, C_θ is the normalization factor and H is real-valued and continuous. As in statistical mechanics, one may regard H as the energy function, θ the temperature and Q a fixed measure in the state space $\mathcal{H}$ (Khinchin [7]).

The questions are: When do we have a weak limit P of P_θ? 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 $N = \{x | H(x) = \inf_y H(y)\}$ denote the set of all minimal energy states. Under the assumption

$$Q\{H(x) < a\} > 0 \quad \text{for } a > \inf H(x), \tag{A1}$$

a necessary condition for the tightness of $\{P_\theta\}$ is $N \neq \varnothing$. 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

$$\exists \varepsilon > 0 \text{ s.t. } \{H(x) \leqslant \min H(x) + \varepsilon\} \text{ is compact} \tag{B1}$$

Received by the editors December 15, 1979 and, in revised form, August 4, 1980.

1980 *Mathematics Subject Classification.* Primary 60B10, 28A50; Secondary 28C20.

Key words and phrases. Characteristic function, covariance operator, Gaussian measure, fiber measure, Hilbert space, Laplace's method, weak convergence.

[1] This research was supported by NSF Grant MCS 76-80762 and the National Science Council of the Republic of China.

0002-9939/81/0000-0327/$02.50

is sufficient for the tightness of $\{P_\theta\}$. With assumptions (A1), (B1), $\mathcal{H} = \mathbf{R}^n$ 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 $\overline{\mathcal{R}(B)}$ (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 $N \neq \varnothing$. Equivalently, F is n.d. (nonnegative definite) and $k = F(m)$ for some m.

Theorem 1 gives a sufficient condition

$$(m + \mathcal{N}(F)) \cap \mathcal{R}(B^{1/2}) \neq \varnothing, \tag{A2}$$

for the existence of P, where $\mathcal{N}(F)$ 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 $Q(dxdy) = Q_X(dx)Q_x(dy)$, where Q_X is a Gaussian measure on X with mean 0 and covariance $\pi_X B \pi_X$ and Q_x'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_x is defined to be the fiber measure of Q on the affine subspace $x + Y$. In the present case let $X = \overline{\mathcal{R}(F)}$ and $Y = \mathcal{N}(F)$. Theorem 2 gives a geometric interpretation that P is the same as $Q_{\pi_{\overline{\mathcal{R}(F)}}m}$, which is the fiber measure of Q on $\pi_{\overline{\mathcal{R}(F)}}m + \mathcal{N}(F) = m + \mathcal{N}(F)$. Moreover, P depends on $m + \mathcal{N}(F)$ 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 $Lg(t) = 0$, $t \notin \mathbf{Z}$, 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

$$V_k = \text{column}\big(r_k(k+1), \ldots, r_k^{(p-2)}(k+1)\big)$$

and

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

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_θ w.r.t. Lebesgue measure is proportional to

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

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_θ is Gaussian.

LEMMA 1. *The characteristic function ψ_θ of P_θ is*

$$\psi_\theta(t) = \exp\big(i\langle G_\theta(F/\theta)m, t\rangle - \tfrac{1}{2}\langle G_\theta t, t\rangle\big)$$

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 $V_n = \text{span}\{e_1, \ldots, e_n\}$, π_n = projection onto V_n and $F_n = \pi_n F \pi_n$. Define

$$(dQ_n/dQ)(x) = C_n \exp\big(-\tfrac{1}{2}(\langle F_n x, x\rangle - 2\langle F_n m, x\rangle)\big);$$

then $Q_n \to P_1$ weakly. Hence, the characteristic function ϕ_n of Q_n converges to ψ_1. In fact, for $t \in V_n$

$$\phi_n(t) = \exp\Big(i\big\langle (F_n + B^{-1})^{-1} F_n m, t\big\rangle - \tfrac{1}{2}\big\langle (F_n + B^{-1})^{-1} t, t\big\rangle\Big).$$

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

$$\psi_1(t) = \exp\Big(i\langle B^{1/2}(I + B^{1/2}FB^{1/2})^{-1}B^{1/2}Fm, t\rangle - \tfrac{1}{2}\langle B^{1/2}(I + B^{1/2}FB^{1/2})^{-1}B^{1/2}t, t\rangle\Big). \qquad \square$$

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 π_D is the projection onto $\mathcal{N}(D)$.*

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 bounded by 1,

$$\frac{\theta}{\theta+\lambda} \to \begin{cases} 0 & \text{if } \lambda \neq 0, \\ 1 & \text{if } \lambda = 0, \end{cases} \qquad \text{as } \theta \downarrow 0.$$

Hence,

$$\left(I + \frac{1}{\theta} D\right)^{-1} = \int_{\sigma(D)} \frac{\theta}{\theta + \lambda} E(d\lambda) \to \Delta E(0) = \pi_D$$

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

By Lemma 2, $G_\theta \to B^{1/2}\pi B^{1/2}$ strongly, where π is the projection onto $\mathfrak{N}(B^{1/2}FB^{1/2})$. Obviously G_θ and $B^{1/2}\pi B^{1/2}$ are bounded by the S-operator B. If we can prove that $G_\theta(F/\theta)m$ converges to some $\hat{m}$ strongly, then $P_\theta \to P$ weakly and P is Gaussian with mean $\hat{m}$ and covariance operator $B^{1/2}\pi B^{1/2}$ (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 $m \in \mathfrak{R}(B^{1/2})$ and m_0 with $B^{1/2}(m_0) = m$. Then

$$G_\theta(F/\theta)m = B^{1/2}m_0 - B^{1/2}(I + B^{1/2}(F/\theta)B^{1/2})^{-1}m_0 \to m - B^{1/2}\pi m_0. \tag{2.1}$$

Therefore, we have

THEOREM 1. *If* (A2) *holds, then $P_\theta \to P$ 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 $m + \mathfrak{N}(F)$.*

PROOF. Let $Y = \mathfrak{N}(F)$, $X = Y^\perp = \overline{\mathfrak{R}(F)}$, π_X and π_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 $\pi_Y B^{1/2}\pi_Y - \pi_Y B\pi_X(\pi_X B\pi_X)^{-1}\pi_X B\pi_Y$; see formulae (6) and (7) in Lemma 2 of Krée and Tortrat [**8**]. First let us prove

$$\pi_Y B^{1/2}\pi B^{1/2}\pi_Y = \pi_Y B\pi_Y - \pi_Y B\pi_X(\pi_X B\pi_X)^{-1}\pi_X B\pi_Y. \tag{2.2}$$

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

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

$$\left(I - B^{1/2}\pi_X(\pi_X B\pi_X)^{-1}\pi_X B^{1/2}\right)z = z - B^{1/2}\pi_X FB^{1/2}u = z - B^{1/2}FB^{1/2}(u) = 0.$$

Since π is bounded, (2.2) holds.

Now we have to relate mean $m - B^{1/2}\pi m_0$ 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 $y \in Y$

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

$$\begin{aligned} m - B^{1/2}\pi m_0 &= m - B^{1/2}\left(I - B^{1/2}\pi_X(\pi_X B\pi_X)^{-1}\pi_X B^{1/2}\right)B^{1/2}m \\ &= B\pi_X(\pi_X B\pi_X)^{-1}\pi_X m, \end{aligned}$$

$$\begin{aligned} \langle y, m - B^{1/2}\pi m_0\rangle &= \langle y, B\pi_X(\pi_X B\pi_X)^{-1}\pi_X m\rangle \\ &= \langle y, \pi_Y B\pi_X(\pi_X B\pi_X)^{-1}\pi_X m\rangle \\ &= \langle(\pi_X B\pi_X)^{-1}\pi_X B\pi_Y y, \pi_X m\rangle. \end{aligned}$$

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 $\overline{\mathfrak{R}(B^{1/2}\pi B^{1/2})} \subseteq Y$). $\square$

Let T be a bounded linear operator from $\mathcal{H}$ to $\mathcal{H}$ and $\hat{m}$ be a fixed element in $\mathcal{H}$. The Gaussian measure with mean $\hat{m}$ and covariance TBT^* is called the induced measure of Q by $\hat{m} + T$. 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*

$$\mathfrak{N}(F) \subseteq \mathfrak{R}(B), \tag{A3}$$

$B^{-1/2}\pi B^{1/2}$ *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])

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

lead to the conclusion. $\square$

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 $B(e_n) = e_n/n^2$, $x = \Sigma_1^\infty e_n/n$ and $y = (\Sigma_2^\infty 1/n^2)e_1 + \Sigma_2^\infty (-1/n)e_n$. Then $x \perp y$. 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 $\mathfrak{N}(F) = \operatorname{span}\{B^{1/2}(x)\}$. $y \perp \mathfrak{N}(FB^{1/2})$, $x + y \in \mathfrak{R}(B^{1/2})$. Let $z = x + y$; then $B^{1/2}(z) = z$. $\pi B^{1/2}(z) = x \notin \mathfrak{R}(B^{1/2})$. 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 $\bar{m}$ such that $\{G_\theta(F/\theta)\bar{m}\}$ is unbounded. Otherwise, by the principle of uniform boundedness, there will be a contradiction.

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

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

and $\{P_\theta\}$ is not tight. $\square$

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

1. N. Dunford and J. T. Schwartz, *Linear operators*, vol. II, Interscience, New York, 1963.
2. U. Grenander, *Probabilities on algebraic structures*, Wiley, New York, 1963.
3. ______, *Lectures in pattern theory*, vol. III (to appear).
4. ______, *Solve the second limit problem in metric pattern theory*, Report on Pattern Analysis No. 83, Div. of Appl. Math., Brown University, Providence, R. I., 1979.
5. C. R. Hwang, *Frozen patterns and minimal energy states*, Ph. D. thesis, Brown University, Providence, R. I., 1978.
6. ______, *Laplace's method revisited: Weak convergence of probability measures*, Ann. Probability **8** (1980).
7. A. I. Khinchin, *Mathematical foundation of statistical mechanics*, Dover, New York, 1957.
8. P. Krée and A. Tortrat, *Désintégration d'un loi gaussienne μ dans une somme vectorielle*, C. R. Acad. Sci. Paris Sér. A **277** (1973), 695–697.
9. B. S. Rajput, *The support of Gaussian measures on Banach spaces*, Theor. Probability Appl. **17** (1972), 728–734.
10. F. Riesz and B. Nagy, *Functional analysis*, McGraw-Hill, New York, 1955.

INSTITUTE OF MATHEMATICS, ACADEMIA SINICA, NANKANG, TAIPEI, TAIWAN, REPUBLIC OF CHINA