

# 混合正規分布に対する統計的 maximum 演算に関する研究

## A Study on Statistical Maximum Operation for Gaussian Mixture Model

電気電子情報通信工学専攻 東 大貴

Daiki AZUMA

### 1. 序論

近年、タイミング解析において遅延ばらつきを統計量として扱う統計的静的遅延解析 (S-STA : Statistical Static Timing Analysis) [1,2] や組電池の寿命ばらつき解析[3]など、設計に統計的手法が利用されている。これらの手法では、 $n$  個のばらつき量の最大 (最小) 値を求める必要が生じる。これを正確に計算するには、各ばらつき量の分布が正規分布に従い、互いに相関を持つ場合、確率密度関数 (PDF : Probability Density Function) が  $n!$  個の項を持ち[4]、その計算は容易ではない。このため 2 個のばらつき量の統計的 maximum 演算 (2Max 演算) を  $n-1$  回繰り返して  $n$  個の最大値を求める手法が研究されてきた。このとき、2 つの正規分布の最大値の分布はもはや正規分布ではなく [1,2]、非正規分布の表現手法が必要となる。PDF が複数個の正規分布の線形和で表される混合正規分布 (GMM : Gaussian Mixture Model) は、非正規分布を表現でき、確率変数間の相関も容易に扱えるため有用であり、音声・音響の分野でも用いられている [5,6]。

本研究では、 $n$  個の正規分布ばらつきをする確率変数  $D_d$  ( $1 \leq d \leq n$ ) の最大値  $\mathcal{M}(n) = \text{Max}[D_d | 1 \leq d \leq n]$  の分布を GMM で近似する問題を考える。このためには GMM 同士の演算手法が必要となるが、 $k$  個の成分を持つ (つまり  $k$  個の正規分布で構成される)  $k$ -GMM と  $h$ -GMM の最大値の分布は、 $2kh$  個の成分 (正規分布ではない) を持つ [2]。従ってこの各成分を正規分布で近似すれば、最大値の分布を  $m$ -GMM ( $m = 2kh$ ) で近似できる。よって GMM 同士の演算を繰り返すと、演算結果の成分の個数が  $n$  の指数関数で増加し、計算効率が悪い。このため既存手法では、毎回の 2Max 演算結果を、平均と分散が等しく PDF の違いが最小の 2-GMM で近似する演算手法を提案してきた [2,8] が、近似誤差の蓄積で、 $\mathcal{M}(n)$  を近似する GMM の精度は低下する [8]。

一方、統計的手法では確率変数間の相関が重要である

が、これを効率よく扱うために、ばらつき量を幾つかの互いに独立な説明変数の線形和で表す表現手法がしばしば用いられる [1-3]。この場合、2Max 演算結果を 2-GMM に近似する際、説明変数の感度、すなわち最大値分布と説明変数の共分散を計算する必要がある。しかし、最大値分布を近似する GMM の分散を、元の最大値分布のそれに一致させると、各説明変数の感度 (共分散) を最大値分布のそれに一致させられない場合が生じる [8]。そこで、GMM の PDF を最大値分布のそれに近づけるために、分散を一致させ、感度 (共分散) の精度を犠牲にすると、2Max 演算の繰り返しで確率変数間の相関に誤差が蓄積し、 $\mathcal{M}(n)$  を近似する GMM の精度が落ちる [8]。

本研究では、分散の精度を損なうことなく、共分散の誤差を小さくする感度の決定手法を提案し、最大値  $\mathcal{M}(n)$  の演算精度の観点から、その性能評価を行った。

### 2. 混合正規分布 (GMM)

確率変数  $D$  で表される分布の PDF が  $m$  個の正規分布の PDF の和

$$\varphi(D) = \sum_{i=1}^m P_i \cdot \frac{1}{\sigma_i} \cdot \phi\left(\frac{D-\mu_i}{\sigma_i}\right), \quad \sum_{i=1}^m P_i = 1 \quad (1)$$

であるとき、 $D$  の分布は  $m$ -GMM であると言い、 $m$ -GMM  $D$  と書く。ここで、 $P_i$  は  $D$  の第  $i$  成分  $N(\mu_i, \sigma_i^2)$  ( $1 \leq i \leq m$ ) の混合比で、 $\phi(x)$  は標準正規分布  $N(0,1)$  の PDF である。標準正規分布の累積分布関数 (CDF) は  $\Phi(x)$  で表す。

$m$ -GMM  $D$  の 1 次モーメント  $E[D]$  および 2 次モーメント  $E[D^2]$  はそれぞれ次式で計算できる。

$$E[D] = \sum_{i=1}^m P_i \cdot \mu_i \quad (2)$$

$$E[D^2] = \sum_{i=1}^m P_i \cdot (\sigma_i^2 + \mu_i^2) \quad (3)$$

また、確率変数  $Y \sim N(\mu_Y, \sigma_Y^2)$  と  $D$  の第  $i$  成分との相関係数を  $\rho_i$  とすると、 $D$  と  $Y$  の共分散は次式で得られる。

$$C[D, Y] = \sum_{i=1}^m P_i \cdot \sigma_i \cdot \sigma_Y \cdot \rho_i = \sum_{i=1}^m P_i \cdot C_i[D, Y] \quad (4)$$

以下では、 $D$  の第  $i$  成分と  $Y$  の共分散を  $C_i[D, Y]$  と書く。

m-GMM  $D$  の各成分を表す確率変数を  $D_i$  と書くと、 $D_i$  は、互いに独立な確率変数  $r_g \sim N(0,1)$  ( $1 \leq g \leq G$ ) および  $x_D \sim N(0,1)$  に依存してばらつくものとする。

$$D_i = \mu_i + s_x[D_i] \cdot x_D + \sum_g s_g[D_i] \cdot r_g \quad (5)$$

ここで、 $\mu_i$  は  $D_i$  の平均  $E_i[D] = E[D_i]$  であり、 $s_x[D_i] = C_i[D, x_D]$  および  $s_g[D_i] = C_i[D, r_g]$  は、それぞれ  $D_i$  の局所変数  $x_D$  および共通変数  $r_g$  に対する感度である。このとき、 $D_i$  の分散  $V[D_i]$  は次式となる。

$$V[D_i] = s_x[D_i]^2 + \sum_g s_g[D_i]^2 \quad (6)$$

$m_A$  個の成分を持つ  $m_A$ -GMM  $D_A$  および  $m_B$  個の成分を持つ  $m_B$ -GMM  $D_B$  の PDF がそれぞれ次式で表され、

$$\varphi(D_A) = \sum_{j=1}^{m_A} P_{Aj} \cdot \frac{1}{\sigma_{Aj}} \cdot \phi\left(\frac{D_A - \mu_{Aj}}{\sigma_{Aj}}\right), \quad \sum_{j=1}^{m_A} P_{Aj} = 1 \quad (7)$$

$$\varphi(D_B) = \sum_{k=1}^{m_B} P_{Bk} \cdot \frac{1}{\sigma_{Bk}} \cdot \phi\left(\frac{D_B - \mu_{Bk}}{\sigma_{Bk}}\right), \quad \sum_{k=1}^{m_B} P_{Bk} = 1 \quad (8)$$

各成分  $D_{Aj}$ ,  $D_{Bk}$  ( $1 \leq j \leq m_A, 1 \leq k \leq m_B$ ) が

$$D_{Aj} = \mu_{Aj} + s_x[D_{Aj}] \cdot x_A + \sum_g s_g[D_{Aj}] \cdot r_g \quad (9)$$

$$D_{Bk} = \mu_{Bk} + s_x[D_{Bk}] \cdot x_B + \sum_g s_g[D_{Bk}] \cdot r_g \quad (10)$$

で表されるような  $D_A$ ,  $D_B$  の最大値  $D_M = \text{Max}[D_A, D_B]$  を求めることを考える。最大値  $D_M = \text{Max}[D_A, D_B]$  の分布は、 $D_A$  と  $D_B$  の同時確率密度関数 (JPDF) の第  $j, k$  成分から 2 つの成分  $j, k, 1$  および  $j, k, 2$  が生まれるので、成分数が  $m = 2 \cdot m_A \cdot m_B$  の m-GMM で近似できる[8]。

$$f_m(D_M) = \sum_{j,k,q} P_{j k q} \cdot \frac{1}{\sigma_{j k q}} \cdot \phi\left(\frac{D_M - \mu_{j k q}}{\sigma_{j k q}}\right) \quad (11)$$

こうして得られた m-GMM  $D_M$  の成分数を削減するには、適切な 2 つ (あるいは複数) の成分を選び、それらを 1 次および 2 次モーメントが一致した 1 つの成分で置き換えることにより行う[8]。例えば、選ばれた  $D_M$  の成分が 3 つ組の集合  $U = \{(j, k, q)\}$  で指定されたならば、次式で決定される混合比  $P_{Mi}$ , 平均  $\mu_{Mi}$ , および分散  $\sigma_{Mi}^2$  を持つ成分で置き換える。

$$P_{Mi} = \sum_{(j,k,q) \in U} P_{j k q} \quad (12)$$

$$P_{Mi} \cdot \mu_{Mi} = \sum_{(j,k,q) \in U} P_{j k q} \cdot \mu_{j k q} \quad (13)$$

$$P_{Mi} \cdot (\sigma_{Mi}^2 + \mu_{Mi}^2) = \sum_{(j,k,q) \in U} P_{j k q} \cdot (\sigma_{j k q}^2 + \mu_{j k q}^2) \quad (14)$$

今、こうして成分数を削減して得られた 2-GMM の確率変数を  $M$  と表すと、 $M$  の平均および分散はそれぞれと  $D_M$  のそれらと等しい。しかし、2Max 演算を  $M$  に適用するには、式(9), (10) と同様、 $M$  の第  $i$  成分を表す

確率変数  $M_i$  が次式で表わされている必要がある。

$$M_i = \mu_{Mi} + s_x[M_i] \cdot x_M + \sum_g s_g[M_i] \cdot r_g \quad (15)$$

すなわち、感度  $s_x[M_i]$  と  $s_g[M_i]$  ( $1 \leq g \leq G$ ) を求める必要がある。その際、 $M_i$  の分散  $\sigma_{Mi}^2$  は式(14)で決定されるため、感度は次式を満たす必要がある。

$$\sigma_{Mi}^2 = s_x[M_i]^2 + \sum_g s_g[M_i]^2, \quad (i = 1, 2) \quad (16)$$

同時に、m-GMM  $D_M$  と共通変数  $r_g$  との共分散

$$C[D_M, r_g] = \sum_{j,k} P_{Aj} \cdot P_{Bk} \cdot C_{j k}[D_M, r_g] \quad (17)$$

を、2-GMM  $M$  と  $r_g$  の共分散  $C[M, r_g]$  に一致させたい。

$$C[M, r_g] = \sum_{i \in \{1, 2\}} P_{Mi} \cdot s_g[M_i] = C[D_M, r_g] \quad (18)$$

すなわち各共通変数  $r_g$  の感度  $s_g[M_i]$  は、式(16) と(18)を同時に満たす必要があるが、常にこれが可能とは限らない。そこで本研究では、式(16) と(18) を同時に満たせない場合でも、式(16)は必ず満たすように感度を決定する。そのかわりに  $C[M, r_g]$  の値を、 $C[D_M, r_g]$  の値と一致させることを諦め、理想の値から変化させる。これは分布形状を変化させないためである。ただし共分散の変化は、後の演算、つまり  $M$  と他の変数との演算を行うときに影響をもたらす。なお、感度  $s_x[M_i]$  には式(18)に対応する制約はないので、 $s_g[M_i]$  の決定後、式(16)から決定する。

$$s_x[M_i]^2 = \sigma_{Mi}^2 - \sum_g s_g[M_i]^2 \geq 0 \quad (19)$$

### 3. 共通変数の感度決定手法

本研究では 2 種類の感度決定手法を提案した (感度決定手法 1, 2 と書く) が、ここでは手法 2 を詳しく記す。

まず、式(16) と (18) を同時に満たす感度  $s_g[M_i]$  が存在するかを調べる。次式を満たす実数  $K$  を見出すことができれば、式(18) を満たす感度が得られる。

$$\begin{cases} s_g[M_1] = s_g[D_M] + K \cdot P_{M2} \\ s_g[M_2] = s_g[D_M] - K \cdot P_{M1} \end{cases} \quad (20)$$

従って、式(16)より次式を得るから、

$$\begin{cases} \sigma_{M1}^2 \geq \sum_g (s_g[D_M]^2 + 2 \cdot P_{M2} \cdot s_g[D_M] \cdot K + P_{M2}^2 \cdot K^2) \\ \sigma_{M2}^2 \geq \sum_g (s_g[D_M]^2 - 2 \cdot P_{M1} \cdot s_g[D_M] \cdot K + P_{M1}^2 \cdot K^2) \end{cases} \quad (21)$$

式(21) を満たす実数  $K$  が存在すれば、式(16), (18) を同時に満たす感度が存在する。このような実数  $K$  が存在するか否かは、 $i = 1, 2$  に対して

$$\begin{aligned} \text{DSC}_i &= (\sum_g s_g[D_M])^2 + G \cdot (\sigma_{Mi}^2 - \sum_g s_g[D_M]^2) \geq 0 \quad (22) \end{aligned}$$

が成立し (G は共通変数の個数), かつ次式が成立するかを調べればよい.

$$-P_{M1} \cdot \sqrt{DSC_1} - P_{M2} \cdot \sqrt{DSC_2} \leq \sum_g C[D_M, r_g] \leq P_{M1} \cdot \sqrt{DSC_1} + P_{M2} \cdot \sqrt{DSC_2} \quad (23)$$

これらが成立するならば, 式(21) の不等式から実数 K のとりうる範囲が求まるので, その範囲内の適切な K を定めれば, 式(20) を用いて感度  $s_g[M_i]$  を決定できる.

式(22), (23) が同時に成立しないならば, 式(16)と(18)を同時に満たす感度は存在しないので, 式(18) の共分散  $s_g[D_M] = C[D_M, r_g]$  を変更する.

ここで, M の 2 つの成分の内, 分散が小さい方を h, 大きい方を H と書き (すなわち,  $\sigma_{Mh} \leq \sigma_{MH}$  である), 第 h 成分に対して, 補正係数  $\gamma_h$  を

$$\gamma_h = \frac{\sigma_{Mh}^2}{\sum_g s_g[D_M]^2} \quad (24)$$

とする. これを用いて, 各感度  $s_g[M_h]$  ( $1 \leq g \leq G$ ) を

$$s_g[M_h] = \gamma_h \cdot s_g[D_M] \quad (25)$$

とすると, 式(16) が第 h 成分に対して成り立つ.

第 H 成分の感度  $s_g[M_H]$  は, 式(18) を用いて,

$$s_g[M_H] = \frac{s_g[D_M] - P_{Mh} \cdot s_g[M_h]}{P_{MH}} = \frac{1 - P_{Mh} \gamma_h}{1 - P_{Mh}} \cdot s_g[D_M] \quad (26)$$

とし, これが第 H 成分に対する式(16) を満たすならば, これを用いる. 式(16) が成り立たない場合, 以下の補正係数  $\gamma_H$

$$\gamma_H = \frac{\sigma_{MH}^2}{\sum_g s_g[D_M]^2} \quad (27)$$

を用いて, 感度  $s_g[M_H]$  を

$$s_g[M_H] = \gamma_H \cdot s_g[D_M] \quad (28)$$

とする. これにより, 式(16) は第 H 成分でも成立する. このとき共分散  $s_g[D_M] = C[D_M, r_g]$  は,  $P_{Mh} \cdot \gamma_h + P_{MH} \cdot \gamma_H$  だけ小さい値に変更されたことになる.

以上が感度決定手法 2 の概要である. 感度決定手法 1 に関して簡単に述べると, 手法 2 に比べて共分散  $C[D_M, r_g]$  の変更量が大きいという特徴がある. 手法 1 はその代わり, 2-GMM の各成分の感度が, 2-GMM 全体の感度中に占める割合にも注意を払うものである.

#### 4. 評価実験

提案手法を評価するため, 次式の正規分布を生成し,

1,000 万回のモンテカルロシミュレーションで求めた  $\mathcal{M}(n)$  の分布と比較する.

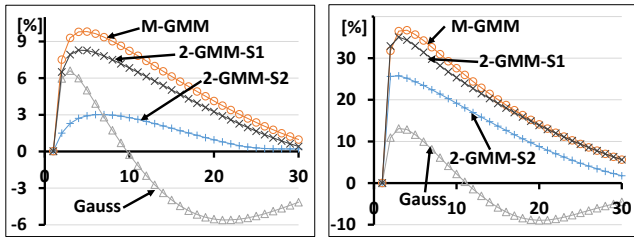
$$D_d = \mu_d + s_x[D_d] \cdot x_d + s_g[D_d] \cdot r \quad (29)$$

比較する手法は, 以下の 4 つである.

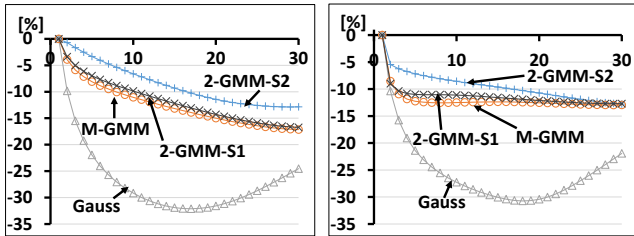
- 1 : Gauss : 正規分布同士の 2Max 演算を行う毎に, 正規分布近似を行うもの.
- 2 : 2-GMM-S1 : 2-GMM 同士の 2Max 演算を行う毎に, 2-GMM 化を行う. 感度決定手法 1 を用いる.
- 3 : 2-GMM-S2 : 2-GMM 同士の 2Max 演算を行う毎に, 2-GMM 化を行う. 感度決定手法 2 を用いる.
- 4 : M-GMM : GMM 同士の 2Max 演算を行う度に 2-GMM 化せず, 成分の個数を増やし続けるもの.

ここでは,  $D_1$  の標準偏差だけを 3 ( $\sigma_1 = 3$ ), それ以外の  $D_d$  の標準偏差を全て 1, 各  $D_d$  の平均  $\mu_d$  は  $\mu_1 = \mu_2 = 0$ ,  $d \geq 3$  で  $\mu_d = (d - 2) \cdot 0.08$  とした場合の結果を紹介する. 相関は全て互いに同じで, 相関係数  $\rho$  を 0.5, 0.9 と変え,  $n = 31$  まで調べた. この場合,  $\text{Max}[D_1, D_2]$  の分布の歪みが大きいため, GMM の有用性が現れる. 図 1 に, 演算回数に対する各  $\mathcal{M}(n)$  の平均, 標準偏差および共分散  $C[\mathcal{M}(n), r]$  の誤差率を示す. 図において, 左側は相関係数  $\rho$  が 0.5, 右側が 0.9 の場合である. 図 1(a) より,  $\rho = 0.9$  のとき, GMM を用いる手法がどれも 2 回目の演算 ( $\mathcal{M}(3)$  の分布との比較) で平均の誤差率が 20% を超えていることが分かる. これは, 1 回目の演算において, 補正係数を用いて共分散の変更が行われ, 共分散の誤差が大きくなったためである. 手法 2-GMM-S2 は共分散の変更量を小さくできたため, 他の GMM を用いた手法より誤差が小さい. また手法 Gauss は, 共分散と分散の間の制約がなく, 共分散の変更が必要ないため, 平均の誤差率は GMM を用いる手法より小さい. しかし  $\mu + \sigma$  の値の誤差は GMM より大きい (PDF の形状の差も大きい).

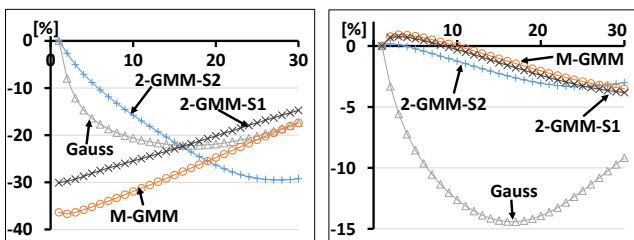
続いて式(29)の正規分布を, 乱数を用いて生成し, 10 万回のモンテカルロシミュレーションで求めた  $\mathcal{M}(n)$  の分布と比較する. このとき,  $D_d$  の平均  $\mu_d$  は 18 から 22 まで, 標準偏差  $\sigma_d$  は 1.26 から 4.4 までの値をとるように乱数を生成した. 各  $D_d$  間の相関は乱数生成の制約上, 小さい値をとりやすいが, 0 から 1 までをとりうる. 今回は  $n=100$  まで調べたが, これを 1,000 回実行した. 表 1 に, 各  $\mathcal{M}(n)$  の平均  $|\mu|$ , 標準偏差  $|\sigma|$  および  $|\mu + \sigma|$  の誤差率を, 1,000 回の平均を取った値で示す. なおこの実験では手法



(a) 平均  $\mu$  の誤差率[%]



(b) 標準偏差  $\sigma$  の誤差率[%]



(c) 共分散  $C[\mu(n),r]$  の誤差率[%]

図 1. 演算回数に対する誤差率 (左  $\rho = 0.5$ , 右  $\rho = 0.9$ )

2-GMM-S2, 2-GMM-S1, Gauss の比較を行った。表を見ると、図 1 の結果と同様、標準偏差の誤差率は手法 2-GMM-S2 の値が常に小さく、優位であった。平均の誤差率も前と同様、手法 Gauss の値が最も小さかった。しかし  $\mu + \sigma$  の値を見ると手法 Gauss が最も悪く、2-GMM-S2 が優位であった。

表 1. 各手法における  $\mu(n)$  の平均誤差率[%]

n	2-GMM-S2			2-GMM-S1			Gauss		
	$\mu$	$\sigma$	$\mu + \sigma$	$\mu$	$\sigma$	$\mu + \sigma$	$\mu$	$\sigma$	$\mu + \sigma$
3	0.0977	0.923	0.0829	0.0929	1.02	0.0890	0.0617	1.75	0.156
5	0.122	1.72	0.144	0.216	2.62	0.174	0.109	3.94	0.307
10	0.155	2.94	0.219	0.300	4.64	0.267	0.124	6.59	0.517
20	0.178	4.11	0.293	0.297	6.00	0.340	0.111	8.09	0.668
30	0.189	4.75	0.324	0.282	6.52	0.372	0.113	8.60	0.717
40	0.200	5.28	0.339	0.274	6.91	0.389	0.109	8.87	0.735
50	0.202	5.70	0.359	0.259	7.25	0.410	0.112	9.06	0.753
60	0.205	5.92	0.364	0.252	7.44	0.417	0.113	9.12	0.754
70	0.198	6.13	0.377	0.237	7.60	0.430	0.115	9.19	0.761
80	0.191	6.20	0.387	0.226	7.67	0.439	0.122	9.16	0.760
90	0.186	6.28	0.391	0.216	7.71	0.443	0.120	9.14	0.753
100	0.181	6.33	0.400	0.204	7.72	0.450	0.125	9.08	0.754

## 5. むすび

本研究では、 $n$  個のばらつき量の最大値  $\mu(n)$  の分布を、 $n-1$  回の 2Max 演算を繰り返して計算する際、各ばらつきが幾つかの説明変数の線形結合で表される場合に必要

となる共分散 (説明変数の感度) の計算手法について考察した。実際の最大値分布を GMM で近似する際、分散と共分散を同時に正しい値にできない場合が生じる。近似した分布を最大値分布に近づけるため、分散を一致させると、共分散の値に誤差が生じ、これが後の 2Max 演算に影響し、 $\mu(n)$  の近似で誤差を生む。感度決定手法 2 は、分散を一致させたまま、共分散の誤差を最小化するものである。この手法 2 の性能を評価する幾つかの実験を行い、 $\mu(n)$  を演算する上での有効性を示した。また、最大値演算を実行する度に 2-GMM 近似を行う従来手法 [2,8] の妥当性も示した。今後の課題として、2-GMM の各成分の感度を決定する際の定数  $K$  の適切な決定方法が挙げられる。また、3 個以上の最大値の分布を 2Max 演算の繰り返しではなく、一気に求め、それを GMM で近似する手法の確立も重要である。

## 参考文献

- [1] D.Blaauw, K.Chopra, A.Srivastave, L.Scheffer, "Statistical timing analysis: From basic principles to state of the art," IEEE Trans. CAD/ICAS, vol.27, no.4, pp.589-607, 2008.
- [2] S.Takahashi and S.Tsukiyama, "A new statistical timing analysis using Gaussian mixture models for delay and slew propagated together," IEICE Trans. Fundamentals, vol.E92-A, no.3, pp.900-911, 2009.
- [3] D.Sasaki, S.Tsukiyama, M.Matsunaga, O.Ishibashi, S. Takahashi, "A statistical method for analyzing lifetime of a battery pack," 2015 IEEE Power & Energy Society General Meeting, paper no. PESGM2015-001311, 2015.
- [4] R.B.Arellano-Valle, M.G.Genton, "On the exact distribution of linear combinations of order statistics from dependent random variables," J. Multivariate Analysis, vol.98, pp.1876-1894, 2007.
- [5] 加藤裕介, 鈴木基之, 伊藤彰則, 牧野正三, "複数の雑音重畳モデルの合成による複数雑音環境に頑健な音響モデルの検討", 情報処理学会研究報告, vol.2004-SLP-51, no.57, pp.1-6, 2004.
- [6] 鈴木基之, 加藤裕介, 伊藤彰則, 牧野正三, "SN 比に頑健なマルチミクスチャー-HMM の性能評価", 電子情報通信学会技術研究報告 SP 2005-31, vol.105, no.133, pp.25-30, 2005.
- [7] D.Sinha, H.Zhou, N.V.Shenoy, "Advances in computation of the maximum of a set of Gaussian random variables," IEEE Trans. CAD/ICAS, vol.26, no.8, pp.1522-1533, 2007.
- [8] N.Yokoyama, D.Azuma, S.Tsukiyama, M.Fukui, "A new algorithm for reducing components of a Gaussian mixture model," IEICE Trans. Fundamentals, vol.E99-A, no.12, pp.2425-2434, 2016.