

無限次元調和解析に基づく音場補間とそのカーネル回帰としての解釈*

©植野夏樹, 小山翔一, 猿渡洋 (東京大学 大学院情報理工学系研究科)

1 はじめに

ある音源を含まない空間内において, 任意に分散配置された複数のマイクロフォンを用い, それらの観測値から音圧の連続的な分布を補間, 推定する問題を扱う. これまで, 特に音場再現のための收音手法として, 平面や球面といった単純な形状のマイクロフォンアレイを用いて, 音場の境界条件を観測することに基づく手法が提案されており, このような問題に対しても適用可能である [1-7]. しかしながら, Fig. 1 に示すような分散マイクロフォンアレイに対してこれらの手法を直接適用することは困難である.

本稿では, 離散的な音圧の観測値が与えられた下での, 斉次ヘルムホルツ方程式の解空間における関数補間として最適化問題を定式化する. さらに, 無限次元調和解析に基づき, 最適解, すなわち補間音場を導出する. また, 機械学習の分野で広く補間問題に用いられるカーネル回帰 [8] との関係を議論し, 提案手法が 0 次球ベッセル関数をカーネル関数として用いた場合のカーネルリッジ回帰と一致することを示す.

2 音場の球面調和関数展開

本節では, 提案手法の理論的背景となる音場の球面調和展開について簡潔に述べる. 3次元空間内のある着目する領域 $\Omega \subseteq \mathbb{R}^3$ に音源が存在しない場合, 領域内の任意の位置 $\mathbf{r} \in \Omega$, 周波数 $\omega \in \mathbb{R}$ における音圧 $u(\mathbf{r}, \omega)$ は, 斉次ヘルムホルツ方程式

$$(\nabla^2 + k^2)u(\mathbf{r}, \omega) = 0 \quad (1)$$

を満たす. ここで, $k = \omega/c$ は波数, c は音速を表す. 以降は単一の周波数についてのみ議論するため, ω は省略し, $u(\mathbf{r})$ のように表記する. 式 (1) の解の集合を H , すなわち

$$H = \{u : \Omega \rightarrow \mathbb{C} \mid (\nabla^2 + k^2)u = 0\} \quad (2)$$

と定める. 式 (1) の任意の解は位置 $\mathbf{r}_0 \in \Omega$ の周りで以下のように球面調和関数展開することが可能で

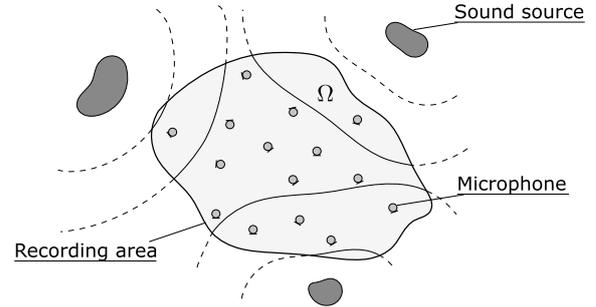


Fig. 1: Sound field interpolation using distributed microphones.

ある.

$$u(\mathbf{r}) = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \hat{u}_{\nu}^{\mu}(\mathbf{r}_0)^* \varphi_{\nu}^{\mu}(\mathbf{r} - \mathbf{r}_0) \quad (3)$$

ここで, $\varphi_{\nu}^{\mu}(\mathbf{r} - \mathbf{r}_0) \in \mathbb{C}$ 及び $\hat{u}_{\nu}^{\mu}(\mathbf{r}_0) \in \mathbb{C}$ はそれぞれ次数 ν , 位数 μ の基底関数及び展開係数であり, 基底関数 $\varphi_{\nu}^{\mu}(\mathbf{r})$ は球面座標表示 $\mathbf{r} = (r, \theta, \phi)$ を用いて,

$$\varphi_{\nu}^{\mu}(\mathbf{r}) = \sqrt{4\pi} j_{\nu}(kr) Y_{\nu}^{\mu}(\theta, \phi). \quad (4)$$

と定義される. ただし, $j_{\nu}(\cdot)$ は ν 次の第一種球ベッセル関数, $Y_{\nu}^{\mu}(\cdot)$ は次数 ν , 位数 μ の球面調和関数である. 式 (3) は無限個の要素に関する総和計算であるが, 以降では有限次元ベクトルの記法と同様に,

$$u(\mathbf{r}) = \boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^{\top} \hat{\mathbf{u}}(\mathbf{r}_0) \quad (5)$$

と表す. ここで, $\boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0) \in \mathbb{C}^{\infty}$ 及び $\hat{\mathbf{u}}(\mathbf{r}_0) \in \mathbb{C}^{\infty}$ は, それぞれ基底関数 $\varphi_{\nu}^{\mu}(\mathbf{r} - \mathbf{r}_0)$ 及び係数 $\hat{u}_{\nu}^{\mu}(\mathbf{r}_0)$ を並べた無限次元ベクトルである.

これらの基底関数及び係数に関して, ベッセル関数の加法定理 [9] より, 任意の $\mathbf{r}, \mathbf{r}_0, \mathbf{r}'_0 \in \Omega$ 及び $u, v \in H$ について以下の式が成立する.

$$\boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^{\top} \boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}'_0)^* = j_0(k\|\mathbf{r}_0 - \mathbf{r}'_0\|) \quad (6)$$

$$\hat{\mathbf{u}}(\mathbf{r}_0)^{\text{H}} \hat{\mathbf{v}}(\mathbf{r}_0) = \hat{\mathbf{u}}(\mathbf{r}'_0)^{\text{H}} \hat{\mathbf{v}}(\mathbf{r}'_0) \quad (7)$$

ここで, $\|\cdot\|$ はユークリッドノルムを表す. したがっ

*Sound field interpolation based on harmonic analysis of infinite order and its interpretation as kernel regression. by UENO, Natsuki, KOYAMA, Shoichi, and SARUWATARI, Hiroshi (Graduate School of Information Science and Technology, The University of Tokyo).

て関数 $\|\cdot\|_H : H \rightarrow [0, \infty]$ を

$$\begin{aligned} \|u\|_H &= \hat{\mathbf{u}}(\mathbf{r}_0)^H \hat{\mathbf{u}}(\mathbf{r}_0) \\ &= \left(\sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} |\hat{u}_{\nu}^{\mu}(\mathbf{r}_0)|^2 \right)^{\frac{1}{2}} \end{aligned} \quad (8)$$

と定義すると、この値は収束、発散も含め \mathbf{r}_0 に依存せず u のみによって定まり、 H における（無限大の値を許す）ノルムの性質を満たす。 u の平面波展開

$$u(r, \theta, \phi) = \int_0^{\pi} \int_0^{2\pi} \tilde{u}(\theta', \phi') e^{ikr \langle (\theta, \phi), (\theta', \phi') \rangle} d\theta' d\phi' \quad (9)$$

が存在するとき、式 (8) のノルムの二乗は、平面波の総パワーの定数倍

$$\frac{1}{4\pi} \int_0^{\pi} \int_0^{2\pi} |\tilde{u}(\theta', \phi')|^2 d\theta' d\phi' \quad (10)$$

に一致する。式 (9) において、 $\langle (\theta, \phi), (\theta', \phi') \rangle$ は 2 つの方向 (θ, ϕ) 及び (θ', ϕ') のなす角の余弦を表す。

3 無限次元調和解析に基づく音場補間

Figure 1 のように M 個の音圧マイクロフォンを用い、 Ω 内の位置 $\mathbf{r}_1, \dots, \mathbf{r}_M$ における音圧 s_1, \dots, s_M を観測したとする。このとき、音圧値の補間問題は以下のように定式化できる。

$$\underset{u \in H}{\text{minimize}} \mathcal{J} = \sum_{m=1}^M |u(\mathbf{r}_m) - s_m|^2 + \lambda \|u\|_H^2 \quad (11)$$

ここで、 $\lambda \in [0, \infty)$ は正則化の強さを表すハイパーパラメータである。第 1 項は観測値との補間誤差を表す損失項、第 2 項は音場のノルムを表す正則化項である。

式 (11) は、 $\mathbf{r}_0 \in \Omega$ を展開中心とする球面調和展開を用いて、

$$\begin{aligned} \mathcal{J} &= \sum_{m=1}^M |\boldsymbol{\varphi}(\mathbf{r}_m - \mathbf{r}_0)^T \hat{\mathbf{u}}(\mathbf{r}_0) - s_m|^2 + \lambda \hat{\mathbf{u}}(\mathbf{r}_0)^H \hat{\mathbf{u}}(\mathbf{r}_0) \\ &= \left(\boldsymbol{\Phi}(\mathbf{r}_0)^T \hat{\mathbf{u}}(\mathbf{r}_0) - \mathbf{s} \right)^H \left(\boldsymbol{\Phi}(\mathbf{r}_0)^T \hat{\mathbf{u}}(\mathbf{r}_0) - \mathbf{s} \right) \\ &\quad + \lambda \hat{\mathbf{u}}(\mathbf{r}_0)^H \hat{\mathbf{u}}(\mathbf{r}_0) \end{aligned} \quad (12)$$

と表せる。ここで、

$$\mathbf{s} = [s_1, \dots, s_M]^T \in \mathbb{C}^M \quad (13)$$

$$\boldsymbol{\Phi}(\mathbf{r}_0) = [\boldsymbol{\varphi}(\mathbf{r}_1 - \mathbf{r}_0) \quad \dots \quad \boldsymbol{\varphi}(\mathbf{r}_M - \mathbf{r}_0)] \in \mathbb{C}^{\infty \times M} \quad (14)$$

である。式 (12) は $\hat{\mathbf{u}}(\mathbf{r}_0)$ に関する二次多項式になっているため、これを最小化する $\hat{\mathbf{u}}(\mathbf{r}_0) = \arg \min_{\hat{\mathbf{u}}(\mathbf{r}_0)} \mathcal{J}$ は、

$$\hat{\mathbf{u}}(\mathbf{r}_0) = \boldsymbol{\Phi}(\mathbf{r}_0)^* (\boldsymbol{\Phi}(\mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0)^* + \lambda \mathbf{I}_M)^{-1} \mathbf{s} \quad (15)$$

で与えられる。ここで \mathbf{I}_M は $M \times M$ の単位行列である。従って、式 (11) の最適解 $\hat{u} = \arg \min_{u \in H} \mathcal{J}$ は、

$$\begin{aligned} \hat{u}(\mathbf{r}) &= \boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^T \hat{\mathbf{u}}(\mathbf{r}_0) \\ &= \boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0)^* (\boldsymbol{\Phi}(\mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0)^* + \lambda \mathbf{I})^{-1} \mathbf{s} \end{aligned} \quad (16)$$

となる。ここで、式 (6) を用いると、

$$\begin{aligned} (\boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0))_m &= \boldsymbol{\varphi}(\mathbf{r} - \mathbf{r}_0)^T \boldsymbol{\varphi}(\mathbf{r}_m - \mathbf{r}_0) \\ &= j_0(k \|\mathbf{r} - \mathbf{r}_m\|) \end{aligned} \quad (17)$$

$$\begin{aligned} (\boldsymbol{\Phi}(\mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0)^*)_{m,m'} &= \boldsymbol{\varphi}(\mathbf{r}_m - \mathbf{r}_0)^T \boldsymbol{\varphi}(\mathbf{r}_{m'} - \mathbf{r}_0) \\ &= j_0(k \|\mathbf{r}_m - \mathbf{r}_{m'}\|) \end{aligned} \quad (18)$$

が成立する。ただし、 $(\cdot)_m$ はベクトルの m 番目の要素、 $(\cdot)_{m,m'}$ は行列の (m, m') 要素を表す。したがって、 $\boldsymbol{\Psi} = \boldsymbol{\Phi}(\mathbf{r}_0)^T \boldsymbol{\Phi}(\mathbf{r}_0)^* \in \mathbb{C}^{M \times M}$ とおくと、式 (16) は、

$$\hat{u}(\mathbf{r}) = \sum_{m=1}^M \left((\boldsymbol{\Psi} + \lambda \mathbf{I})^{-1} \mathbf{s} \right)_m j_0(k \|\mathbf{r} - \mathbf{r}_m\|) \quad (19)$$

と表すことができる。式 (19) は最適な音場の補間 \hat{u} が \mathbf{r}_0 の位置に依存せず、マイクロフォン位置 \mathbf{r}_m とその点における観測音圧 s_m のみによって決まることを意味している。

4 カーネル回帰としての解釈

式 (19) はカーネル回帰 [8] として解釈することも可能である。カーネル回帰とは、入出力のペア $(\mathbf{r}_1, s_1), \dots, (\mathbf{r}_M, s_M)$ が与えられたとき、未知の入力 \mathbf{r} に対する予測値 $\hat{f}(\mathbf{r})$ を、正定値カーネル κ を用いて

$$\hat{f}(\mathbf{r}) = \sum_{m=1}^M \alpha_m \kappa(\mathbf{r}, \mathbf{r}_m) \quad (20)$$

の形で求める手法である。これは以下の最適化問題

$$\underset{f \in H_{\kappa}}{\text{minimize}} L((s_1, f(\mathbf{r}_1)), \dots, (s_M, f(\mathbf{r}_M))) + g(\|f\|_{H_{\kappa}}) \quad (21)$$

の解 \hat{f} が式 (20) の形で表されることを主張する Representer 定理 [10] を根拠にもつ。ここで、 $L(\cdot)$ は任意の損失関数、 $g(\cdot)$ は任意の狭義単調増加関数、 H_{κ} は正定値カーネル κ に対応する再生核 Hilbert 空間、 $\|\cdot\|_{H_{\kappa}}$ は H_{κ} における内積が定めるノルムである。

カーネルリッジ回帰の場合、 $\mathbf{s} = [s_1, \dots, s_M]^T$ として、

$$\hat{f}(\mathbf{r}) = \sum_{m=1}^M ((\mathbf{K} + \lambda \mathbf{I})^{-1} \mathbf{s})_m \kappa(\mathbf{r}, \mathbf{r}_m) \quad (22)$$

となる。ここで、 $\mathbf{K} \in \mathbb{C}^{M \times M}$ は

$$(\mathbf{K})_{m,m'} = \kappa(\mathbf{r}_m, \mathbf{r}_{m'}) \quad (23)$$

で与えられる (半) 正定値行列である。このとき、カーネル関数として

$$\kappa(\mathbf{r}, \mathbf{r}') = j_0(k\|\mathbf{r} - \mathbf{r}'\|) \quad (24)$$

を用いると、式 (23) で与えられる \mathbf{K} は半正定値性を満たし、式 (22) は式 (19) と一致する。すなわち、提案手法は 0 次の球ベッセル関数をカーネルとして用いた場合のカーネルリッジ回帰と等しい。

式 (11) に示す以外の、より一般的な音場補間の定式化に対しても、カーネル回帰と同様に以下の Representer 定理が成立する。

定理 1. 位置 $\mathbf{r}_1, \dots, \mathbf{r}_M \in \Omega$ における音圧の観測値を s_1, \dots, s_M とすると、任意の損失関数 $L(\cdot)$ 及び狭義単調増加関数 $g(\cdot)$ について、斉次ヘルムホルツ方程式の解空間 H における以下の最適化問題

$$\underset{u \in H}{\text{minimize}} \quad L((s_1, u(\mathbf{r}_1)), \dots, (s_M, u(\mathbf{r}_M))) + g(\|u\|_H) \quad (25)$$

の最適解 \hat{u} は

$$\hat{u}(\mathbf{r}) = \sum_{m=1}^M \alpha_m j_0(k\|\mathbf{r} - \mathbf{r}_m\|) \quad (26)$$

の形で表現できる。ここで H 及び $\|\cdot\|_H$ はそれぞれ式 (2) 及び (8) で定義される。

証明 . 付録参照。

5 数値実験

5.1 実験条件

提案手法の有効性を評価するため、数値実験を行った。3次元空間において、原点中心、半径 1.0m の球状領域 Ω に $M = 64$ 個の音圧マイクロフォンを配置した。マイクロフォンの配置は Ω における一様分布に従って決定した。真の音圧値に対し、SN 比 20 dB の正規乱数をノイズとして付加したものを観測信号とし、提案手法によって観測信号から Ω における音場を補間した。ガウスカネル $\kappa(\mathbf{r}, \mathbf{r}') = e^{-(\gamma k\|\mathbf{r} - \mathbf{r}'\|)^2}$ を用いたカーネルリッジ回帰を比較手法とし、補間精度を比較した。ここで、ガウスカネルにおけるパラメータは実験的に $\gamma = 0.5$ と定めた。提案手法の式 (19) 及び比較手法の式 (22) において、いずれも正則化パラメータは $\lambda = 10^{-2}$ とした。

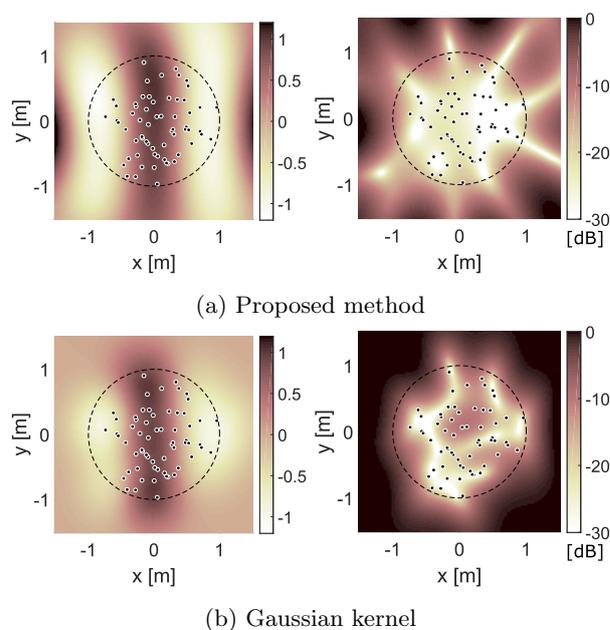


Fig. 2: Interpolated sound pressure distribution (left) and normalized squared error distribution (right) on xy plane at $z = 0$.

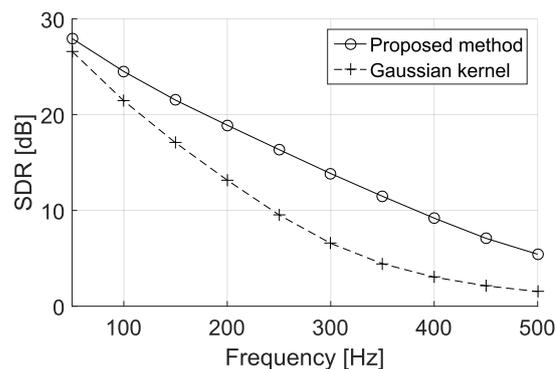


Fig. 3: Relationship between frequency and SDR.

5.2 実験結果

Figure 2 は、真の音場を x 軸の正方向に進行する周波数 200 Hz の平面波としたときの、 xy 平面内の補間音場 (左) 及び正規化二乗誤差 (右) である。黒い中点は $z = 0$ の xy 平面上に射影したマイクロフォン位置を表す。提案手法は、ガウスカネルを用いた補間よりも高い補間精度を達成している。これは後者のカーネルは斉次ヘルムホルツ方程式を満たさず、より一般的な関数空間において補間を行っているのに対し、前者はヘルムホルツ方程式の斉次解の中で補間を行っているためであると考えられる。

さらに、周波数と補間精度の関係を、以下に定義する Signal-to-Distortion Ratio (SDR) の値を用いて定

量的に評価した。

$$\text{SDR}(\omega) = 10 \log_{10} \frac{\int_{\Omega} |u(\mathbf{r}, \omega)|^2 d\mathbf{r}}{\int_{\Omega} |u(\mathbf{r}, \omega) - \hat{u}(\mathbf{r}, \omega)|^2 d\mathbf{r}} \quad (27)$$

ここで、 $u(\mathbf{r}, \omega)$ 及び $\hat{u}(\mathbf{r}, \omega)$ は、それぞれ真の音場及び補間音場である。また、 Ω における積分は、空間を 0.02 m 間隔でサンプリングして計算した。異なる 16 種類のマイクロフォン配置及び、一様分布に従って決定した異なる 64 方向の平面波に対して、それぞれ周波数 50 Hz から 500 Hz まで 50 Hz 毎に SDR を計算し、各周波数における平均値を Fig. 3 にプロットした。いずれの周波数においても、提案手法がガウスカネルを用いた補間よりも高い SDR を達成している。

6 おわりに

無限次元調和解析に基づき、離散的な音圧の観測値から領域全体における連続的な音圧の分布を補間する手法を提案した。カーネル回帰との関係を議論し、提案手法が 0 次球ベッセル関数をカーネルとして用いた場合のカーネルリッジ回帰と一致することを示した。さらに、より一般的な音場補間の定式化における Representer 定理を示した。数値実験によって分散配置マイクロフォンを用いた音場補間を行い、提案手法がガウスカネルを用いた関数補間の手法よりも高い補間精度を実現することを示した。

■謝辞 本研究は、JSPS 科研費 JP15H05312, JP16H01735 及びセコム科学技術振興財団の助成を受けた。

7 付録

定理 1 の証明を以下に示す。

証明 . H における内積を

$$\langle u, v \rangle_H = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \hat{u}_{\nu}^{\mu}(\mathbf{r}_0) \hat{v}_{\nu}^{\mu}(\mathbf{r}_0)^* \quad (28)$$

と定義する。この値が \mathbf{r}_0 に依存しないことは式 (7) から示される。また、関数 $\psi_{\mathbf{r}'} \in H$ を

$$\psi_{\mathbf{r}'}(\mathbf{r}) = j_0(k\|\mathbf{r} - \mathbf{r}'\|) \quad (29)$$

と定義すると、

$$\langle u, \psi_{\mathbf{r}} \rangle_H = \sum_{\nu=0}^{\infty} \sum_{\mu=-\nu}^{\nu} \hat{u}_{\nu}^{\mu}(\mathbf{r}) \delta_{\nu,0} = \hat{u}_0^0(\mathbf{r}) = u(\mathbf{r}) \quad (30)$$

が成立する。すなわち $\psi_{\mathbf{r}}$ は H における再生核である。再生核 $\psi_{\mathbf{r}_1}, \dots, \psi_{\mathbf{r}_M}$ 及び、

$$\langle u_{\perp}, \psi_{\mathbf{r}_m} \rangle_H = 0, \quad \forall m \in \{1, \dots, M\} \quad (31)$$

を満たす u_{\perp} を用いて、 u を

$$u = \sum_{m=1}^M \alpha_m \psi_{\mathbf{r}_m} + u_{\perp} \quad (32)$$

と展開すると、

$$\begin{aligned} u(\mathbf{r}_m) &= \left\langle \sum_{m'=1}^M \alpha_{m'} \psi_{\mathbf{r}_{m'}} + u_{\perp}, \psi_{\mathbf{r}_m} \right\rangle_H \\ &= \sum_{m'=1}^M \alpha_{m'} \langle \psi_{\mathbf{r}_{m'}}, \psi_{\mathbf{r}_m} \rangle_H \end{aligned} \quad (33)$$

より、 $u(\mathbf{r}_m)$ は u_{\perp} に依存しないため、損失関数 L の値も u_{\perp} には依存しない。一方、ノルム罰則項に関しては、任意の u_{\perp} について、 $\sum_{m=1}^M \alpha_m \psi_{\mathbf{r}_m}$ と u_{\perp} の直交性より、

$$\begin{aligned} g(\|u\|_H) &= g\left(\sqrt{\left\| \sum_{m'=1}^M \alpha_{m'} \psi_{\mathbf{r}_{m'}} \right\|_H^2 + \|u_{\perp}\|_H^2}\right) \\ &\geq g\left(\left\| \sum_{m'=1}^M \alpha_{m'} \psi_{\mathbf{r}_{m'}} \right\|_H\right) \end{aligned} \quad (34)$$

が成立する。等号成立は $u_{\perp} = 0$ のときのみであるため、式 (25) の最適解は必ず $u_{\perp} = 0$ 、すなわち式 (26) の形になる。

参考文献

- [1] S. Koyama, et al., *IEEE J. ASLP.*, 21 (4), 685–696, 2013.
- [2] M. A. Poletti, *J. Acoust. Soc. Am.*, 53 (11), 1004–1025, 2005.
- [3] J. Meyer and G. Elko, In *Proc. ICASSP*, II–1781–1784, 2002.
- [4] T. D. Abhayapala and D. B. Ward, In *Proc. ICASSP*, II–1949–1952, 2002.
- [5] B. Rafaely, *IEEE J. ASLP.*, 16 (4), 740–747, 2008.
- [6] S. Koyama, et al., *IEEE/ACM. J. ASLP.*, 22 (10), 1546–1557, 2014.
- [7] S. Koyama, et al., *J. Acoust. Soc. Am.*, 139 (3), 1024–1036, 2016.
- [8] K. P. Murphy, “Machine learning: a probabilistic perspective,” MIT press, 2012.
- [9] P. A. Martin, “Multiple scattering: interaction of time-harmonic waves with N obstacles,” Cambridge University Press, 2006.
- [10] B. Schölkopf, et al., In *Proc. COLT/EuroCOLT, LNAI 2111*, 416–426, 2001.