

SVD フィルタによるカルマンフィルタとニューラルネットワーク併用による磁極形状最適化

Electromagnetic Device Optimization Using Kalman Filter by SVD Filter and Neural Network

○木村昭穂, 松坂知行

Akio Kimura, Tomoyuki Matsuzaka

八戸工業大学

Hachinohe Institute of Technology

キーワード : SVD フィルタ (Singular Value Decomposition), カルマンフィルタ (Kalman filter), ニューラルネットワーク (neural network), 形状最適化 (shape optimization)

連絡先 : 〒031-8501 八戸市大字妙字大開 881-1 八戸工業大学 システム情報工学科

木村昭穂, Tel: (0178) 25-8144, Fax (0178)-25-1691, E-mail: kimura@hi-tech.ac.jp

1. はじめに

電子計算機の高性能化に伴い有限要素法, 境界要素法に代表される数値計算技術は, グラフィカルインタフェースの普及により使いやすくなり, 電気・電子機器の設計や改良に大いに利用されている。機器の設計や改良においては, 所望の性能が備わるようにするためには, 経験をもとに試行錯誤が必要で多くの時間と労力を費やさなければならない場合が多々ある。このような背景のもとに, 電気・電子機器の設計目標値として与えられた電界や磁束分布を満たす形状, 寸法を決定する最

適化手法の応用が精力的に行われている。最適化手法としては, Simulated Annealing 法¹⁾, 遺伝的アルゴリズム²⁾, ニューラルネットワーク (以下 NN と略記)^{3), 4)} 等の提案がなされている。ここでは NN の適用について考察する。ニューラルネットワークの学習方法として代表的ものに誤差逆伝播法 (BP 法) がある。誤差逆伝播法は, 学習回数が非常に大きくなることや, 局所解に落ち込みやすいなどの欠点があるので, 誤差を所定の誤差内に収束させるためには慣性項の追加や, シグモイド関数に代わる伝達関数を使用するな

どのテクニックが用いられている。このようなことからニューラルネットワークでは、より有効な学習法が要求される。学習回数の軽減を図るために、カルマンフィルタ⁵⁾を用いた学習法が報告されている。このカルマンフィルタの計算には、U-D分解フィルタが適用されているが、アルゴリズムが複雑で分かりにくいという指摘⁸⁾がなされている。そこで、アルゴリズムの簡略化を図るためにSVD(Singular Value Decomposition)フィルタの適用を試みた。尚、本計算には、アルゴリズムの数式表現をより忠実に記述できる言語の一つである Matlab を用いた。

2. ニューラルネットワークの構成^{5) 10) ~ 12)}

形状最適化最適化計算に NN を適用するに当たり、カルマンフィルタを利用する。カルマンフィルタ・ニューラルの併用法を図1に示す。NNは入力層と1層以上の中間層および出力層からなる階層構造をしており、教師データにより学習を行う。同図では中間層を1層とした3層構造について示している。図の x_i は入力信号、 w_{ji} はユニット間の重み結合係数、 h_k は出力信号、 y_k は教師信号である。ここで、ニューラルネットワークの N 層のユニット j の内部状態を u_j^N 、 $N-1$ 層のユニットからの入力を h_i^{N-1} 、 $N-1$ 層のユニット i から N 層のユニット j への重み係数を $w_{ji}^{N,N-1}$ とすると、ユニット j の内部状態 u_j^N と出力 h_j^N は、次式のように与えられる。

$$u_j^N = \sum_{i=1}^m w_{ji}^{N,N-1} h_i^{N-1} \quad (1)$$

$$\text{ここで } h_j^N = f(u_j^N) \quad (2)$$

ただし、 m はユニット数、 N は階層、一層のユニットの出力値は h_i^1 、入力信号は x_i である。ネットワークの出力信号 h_k と教師信号 y_k の誤差が指定誤差より大きい場合は、カルマンフィルタにて重み係数 w_{ji} を予測し、同様にネットワークの計算を行い、誤差が指定

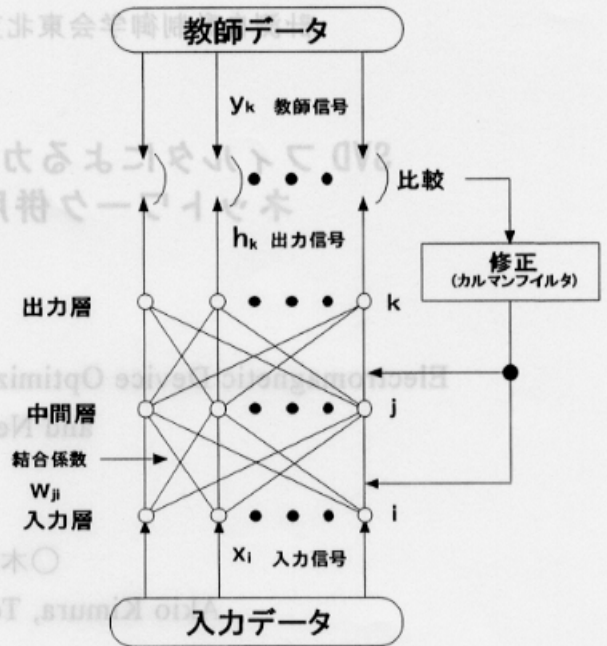


図1 ニューラルネットワークの基本構成

誤差以上或いは、反復回数が指定回数以下ならば計算を繰り返す。 f は伝達関数である。

NNの1回の学習における誤差は、全パターンの出力について教師データとの2乗平均誤差を取るものとする、2乗平均学習誤差 E は

$$E = \sum_{p=1}^{N_p} E_p / N_p \quad (3)$$

$$\text{ここで } E_p = \frac{1}{2} \sum_{k=1}^{N_o} (y_{pk} - h_{pk})^2 \quad (4)$$

で表される。ただし、 N_p は学習パターン数、 N_o は出力層のユニット数、 y_{pk} は教師データ、 h_{pk} は NN の出力である。

3. 拡張カルマンフィルタによる学習

非線形確率システムは線形確率システムに移行されれば、外部入力を有するシステムとして取り扱うことによって、拡張カルマンフ

フィルタを得る事ができる。そこで、カルマンフィルタ⁵⁾⁻⁹⁾を併用したニューラルネットワークの学習法では、推定パラメータに相当する状態ベクトルは、時間的遷移構造を持たないので、時間軸について一定とする。またシステムに誤差の混入がないという条件を付け、システム行列に単位行列を選ぶと次式のような非線形システムの状態推定問題となる。

1) 状態方程式

$$w_{t+1} = F_t w_t + G_t d_t \quad (5)$$

2) 観測方程式

$$y_t = h_t + v_t \quad (6)$$

ここで、 w_t は状態ベクトル(結合重み変数ベクトル)、 y_t は観測ベクトル(NNの教師ベクトル)、 F_t は状態遷移行列、 G_t はシステム駆動行列、 d_t はシステム雑音、 h_t はシステムの観測行列であり¹³⁾、 $h_t = h_t(w_t)$ である。 $h_t(w_t)$ はNNの構造と各ユニットの非線形特性で定まる関数で時不変であるが w_t の関数である。ここで t を反復回数と読みかえる。また、 $t=0$ のとき w_t は w_0 であり、この w_0 は初期値である。ここで w_0 、 d_t はガウス性を満たすものとし、システム雑音がないものとして d_t は零とする。 v_t はガウス白色観測ノイズであり、次式を満たすものである。

$$\begin{aligned} E[v_t] &= 0 \\ E[v_t v_t^T] &= R_t \quad (R_t > 0 \text{ (正定値性)}) \end{aligned} \quad (7)$$

$E[\cdot]$ は期待値である。 R_t は単位行列とし、ガウス白色観測ノイズは非常に小さいものとして無視している。

基本的にカルマンフィルタは、時刻 $t+k$ における状態量 w_{t+k} の最小分散推定値を最小にする推定値 $\hat{w}_{t+k|t}$ を与えるフィルタの設計をすることである。推定値 $\hat{w}_{t+k|t}$ は、時刻 t における時刻 $t+k$ の推定値を表すものとする。 k が $k \geq 1$ のときは最適予測問題となり、 $k=0$ のときは最適濾波問題となる。また $k < 0$ のときは平滑問題となる。カルマンフィ

ルタとニューラルネットワークへの適用は、予測と濾波機能を利用するものである。

式(11)の観測式は状態量 w_t については線形になっているが、推定値が非線形であるので、 $h_t(w_t)$ を推定値のまわりにテイラー展開を行い、高次の項を無視して線形化すると次式が得られる。

$$h_t(w_t) = h_t(\hat{w}_{t|t-1}) + H_t [w_t - \hat{w}_{t|t-1}] \quad (8)$$

ただし、 H_t は観測行列と呼ばれ、この成分は次式のように表される。

$$\frac{\partial h_t}{\partial w_t} \Big|_{w_t = \hat{w}_{t|t-1}} \quad (9)$$

式(8)を式(6)に代入し整理すると、次式が得られる。

$$y_t = h_t(\hat{w}_{t|t-1}) + H_t [w_t - \hat{w}_{t|t-1}] + v_t \quad (10)$$

更に式(10)の一部分を、

$$m_t = y_t - h_t(\hat{w}_{t|t-1}) + H_t \hat{w}_{t|t-1} \quad (11)$$

とし、式(10)を式(11)式を用いて整理する。

式(5)は、次式のように表される。

$$w_{t+1} = F_t w_t \quad (12)$$

$$m_t = H_t w_t + v_t \quad (13)$$

これは外部入力を含まないシステムである。この線形システムに観測更新アルゴリズムであるカルマンフィルタ^{5),8),9)}を適用することにより、次の拡張カルマンフィルタの式を得る。

(a) フィルタ方程式 $\hat{w}_{t|t}$

$$\hat{w}_{t|t} = \hat{w}_{t|t-1} + K_t [y_t - h_t(\hat{w}_{t|t-1})] \quad (14)$$

$$\hat{w}_{t+1|t} = \hat{w}_{t|t} \quad (15)$$

(b) カルマンゲイン K_t

$$K_t = P_{t|t-1} H_t^T [H_t P_{t|t-1} H_t^T + R_t]^{-1} \quad (16)$$

(c) 推定誤差共分散行列 $P_{t|t}$

$$P_{t|t} = P_{t|t-1} - K_t H_t P_{t|t-1} \quad (17)$$

$$P_{t+1|t} = P_{t|t} \quad (18)$$

$$t = 0, 1, 2, 3, \dots$$

(d) 初期条件

$$\hat{w}_{0|-1} = w_0, \quad P_{0|-1} = \beta I \quad (20)$$

$$P_{t+1|t} = \beta \varepsilon I$$

β : 定数, ε : 平均誤差, I : 単位行列

U・D フィルタによるアルゴリズムを以下に述べる。

(14)式において y_t を学習用教師データ, $h_t(\hat{w}_t)$ をネットワーク出力データ, 時間 t を反復回数として重み係数 $\{\hat{w}_t\}$ を推定する。フィルタの計算は, 推定誤差共分散行列 $P_{t,t}$ を U・D 分解して行う。アルゴリズムは以下のようなになる。

- Step 0 :** 初期条件の指定, 式(20)の w_0 は 0 ~ 1 間の乱数, $p_{0,-1}$ は便宜上式(20)で与える, $t = 0$
- Step 1 :** 最初の学習, 教師データ (x_0, y_0) を用意
- Step 2 :** ネットワークの内部状態 u_j および出力 $h_t(\hat{w}_t)$ の計算
- Step 3 :** 式 (9) による観測行列 H_t の計算
- Step 4 :** 式 (16) によるカルマンゲイン K_t の計算
- Step 5 :** 式(14), (17)により \hat{w}_t および p_t を計算し, 観測更新
- Step 6 :** もし, $t <$ ネットワークの計算総数より小さいならば, 式(15), (18) および t を更新 ($t = t + 1$) して, **Step 2** へ
- Step 7 :** もし, 誤差が指定値以下あるいは, 指定回数以上ならば終了
- Step 8 :** 式 (20) により \hat{w}_t および p_t を更新し, $t = 0$ として **Step 1** へ

4. SVD フィルタ

共分散行列の観測更新アルゴリズムは

$$\hat{P} = \bar{P} - \bar{P}H^T [H\bar{P}H^T + R]^{-1} H\bar{P} \quad (21)$$

と与えられ, この逆行列は

$$\hat{P}^{-1} = \bar{P}^{-1} - H^T R^{-1} H \quad (22)$$

のようになる。ただし, 共分散行列 \hat{P} , \bar{P} は

共に非負定値対称である。この SVD を

$$\hat{P}^{-1} = \hat{U}\hat{D}^{-2}\hat{U}^T, \quad \bar{P}^{-1} = \bar{U}\bar{D}^{-2}\bar{U}^T$$

とし, また R^{-1} のコレスキー分解を $R^{-1} = LL^T$ とおくと式(22)は

$$\begin{aligned} \hat{U}\hat{D}^{-2}\hat{U}^T &= \bar{U}\bar{D}^{-2}\bar{U}^T + H^T LL^T H \\ &= \bar{U}(\bar{D}^{-2} + \bar{U}^T H^T LL^T H \bar{U})\bar{U}^T \end{aligned} \quad (23)$$

のように表される。ここで, $[\bar{D}^{-1} \quad \bar{U}^T H^T L]$ の SVD を

$$[\bar{D}^{-1} \quad \bar{U}^T H^T L] = U[D \quad 0]V^T \quad (24)$$

のように表すと,

$$\bar{D}^{-2} + \bar{U}^T H^T LL^T H \bar{U} = UD^2U^T \quad (25)$$

が成立する。これより式(23)は

$$\hat{U}\hat{D}^{-2}\hat{U}^T = \bar{U}UD^2U^T\bar{U}^T \quad (26)$$

となる。上式の両辺を比較すると

$$\hat{U} = \bar{U}U, \quad \hat{D} = D^{-1} \quad (27)$$

を得る。

SVD フィルタルゴリズムを以下に述べる。

Step 0 : 初期値設定

$$\bar{w} = \hat{w}_{t/t-1}, \quad \bar{P} := P_{t/t-1} = \bar{U}\bar{D}^2\bar{U}^T.$$

Step 1 : 観測更新アルゴリズム SVD

$$[\bar{D}^{-1} \quad \bar{U}^T H^T L] = U[D \quad 0]V^T$$

を計算し, 次のようにおく。

$$\hat{U} = \bar{U}U, \quad \hat{D} = D^{-1}$$

$$K := \hat{P}H^T LL^T$$

$$\hat{w} := \bar{x} + K(y - H\bar{w})$$

Step 2 : もし, $t <$ ネットワークの計算総数より小さいならば **Step 1** へ

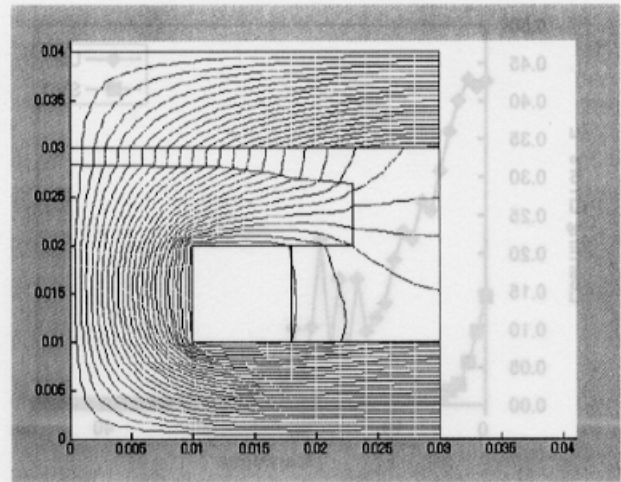
Step 3 : もし, 誤差が指定値以下あるいは, 指定回数以上ならば終了

5. 解析モデルと学習データの作成

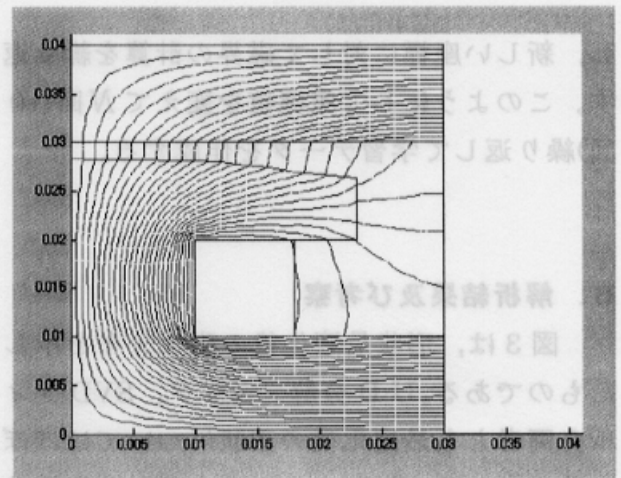
回転機のギャップ中の磁束分布が, 正弦波分布をするような磁極形状の最適化を行う。図 5-3 に, 解析対象である同期電動機の磁極を (回転機モデル A) 示す。コイルのアンペアターンは 2.01×10^3 [AT] とした。固定子のスロットは無視した。何故ならば, ギャップ

中の磁束分布は回転子の形状によって大きく左右され、回転機のギャップ中の磁束分布は固定子スロットによる影響をあまり受けないからである。図のA-B上の磁束密度が正弦分布をするものとし、正弦分布の最大磁束密度は0.5[T]で、A点は90度、C点は0度の位置に対応させる。また、図の境界点 a_1, a_2, \dots, a_{11} は等間隔の点であり、学習データ作成のための考察点である。そして、図の p_1, p_2, \dots, p_{11} は等間隔に配置され、垂直方向にのみ移動するものとし、回転子の直径は $2.8 \times 10^{-2}[m]$ を超えないものとする。さらに、図の境界 $\alpha-\beta$, $\beta-\gamma$, $\delta-\alpha$ は固定境界条件とし、ベクトルポテンシャル A を零とする。また、境界 $\beta-\gamma$ は自然境界である。

磁極形状最適化問題にニューラルネットワークを適用するためには、磁界分布を詳細に計算し、学習データを作成する必要がある。磁界分布の詳細は有限要素法(2D)により行い、学習データ作成のための考察点 a_i ($i=1, \dots, 11$)、移動点 P_i ($i=1, \dots, 11$)での磁束分布を求める。次に、移動点 P_i ($i=1, \dots, 11$)の座標を Δy だけ移動



(a)U-D分解フィルタによる磁束分布



(b)SVDフィルタによる磁束分布

図3 磁束分布

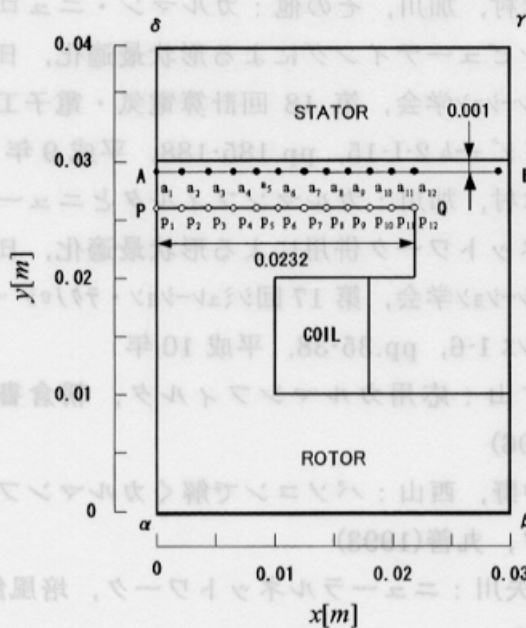


図2 磁極モデル

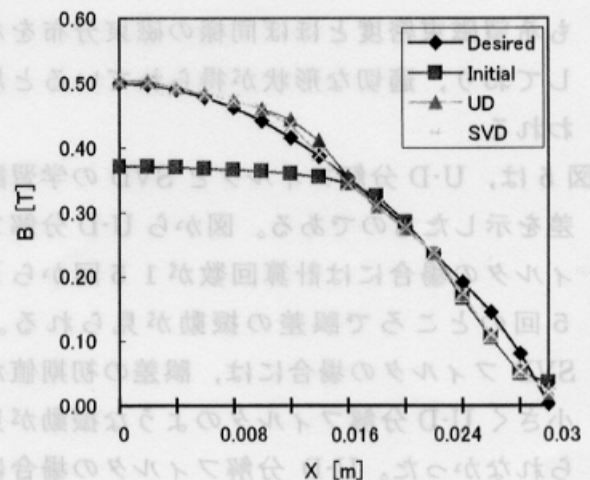


図4 A-B間の磁束密度分布

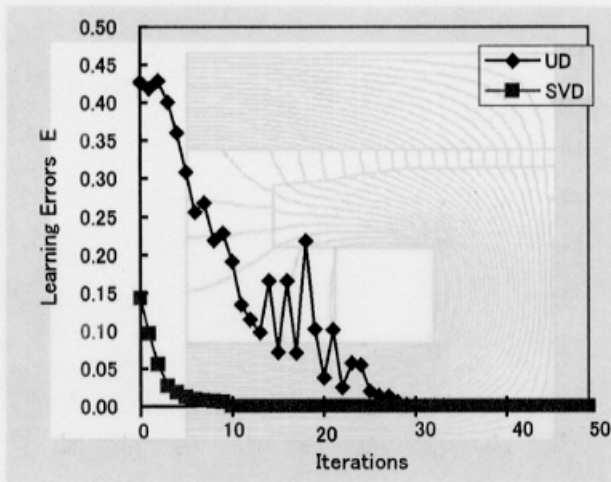


図5 誤差の比較

し、新しい座標に対して磁界の計算を繰り返す。このようにして座標値を変えて N 回(10回)繰り返して学習データを作成する。

6. 解析結果及び考察

図3は、形状最適化後の磁束分布を示したものである。U-D分解フィルタ、SVDフィルタ両者とも最適化後の形状についてはほぼ一致しており、SVDフィルタを用いた計算結果が妥当であることが分かる。

図4は、U-D分解フィルタとSVDフィルタの考察領域A-B間の磁束密度分布を示したものである。U-D分解フィルタ、SVDとも希望磁束密度とほぼ同様の磁束分布をなしており、適切な形状が得られていると思われる。

図5は、U-D分解フィルタとSVDの学習誤差を示したものである。図からU-D分解フィルタの場合には計算回数が15回から25回のところで誤差の振動が見られる。SVDフィルタの場合には、誤差の初期値が小さくU-D分解フィルタのような振動が見られなかった。U-D分解フィルタの場合に計算回数が40回で収束に至ったが、SVDフィルタの場合には収束しなかった。

SVDフィルタの場合U-D分解フィルタと比較するとアルゴリズムが簡単であるが、SVD変換に大きな配列を必要とするために計算に時間が掛かることが分かった。

参考文献

- 1) 高橋, 中田, その他: 最適化問題解析比較用モデル(磁気プレスモデル)の解析及び実験, 日本シミュレーション学会, 第16回計算電気・電子工学シンポジウムII-2, pp.147-150, 平成7年
- 2) 若尾, 小貫, 小川: 三次元磁気シールド問題に対する位相最適化の検討, 電気学会, pp501-506, 1997
- 3) 大塚, 加藤, その他: ニューラルネットワークを用いた電界最適化技術の開発, 電気学会, pp1387-1397, 19995
- 4) S. Ratnajeevan, H. Hoole: Artificial Neural Networks in the Solution of Inverse Electromagnetic Field Problems, IEEE, Tras. Magn., Vol.29, No.2, pp. 1931-1933..
- 5) 村瀬他: カルマン・ニューロコンピューティング, 森北出版(1994)
- 6) 木村, 加川, その他: カルマン・ニューロコンピューティングによる形状最適化, 日本シミュレーション学会, 第18回計算電気・電子工学シンポジウム2-I-15, pp.185-188, 平成9年
- 7) 木村, 加川: カルマンフィルタとニューラルネットワーク併用による形状最適化, 日本シミュレーション学会, 第17回シミュレーション・テクノロジー・コンファレンス1-6, pp.35-38, 平成10年
- 8) 片山: 応用カルマンフィルタ, 朝倉書店(1996)
- 9) 中野, 西山: パソコンで解くカルマンフィルタ, 丸善(1993)
- 10) 矢川: ニューラルネットワーク, 培風館, 1995
- 11) 市川: 階層型ニューラルネットワーク, 共

立出版 (1993)

12)西川, 北村:ニューラルネットワークと計測制御, 朝倉書店 (1995)

13)相良, 秋月他:システム同定, 計測自動制御学会, 昭和 56 年