

派生証券の価格評価における人工知能の活用とその展望

舟橋 秀治

An AI Application To Derivatives Pricing And Its Prospects

Hideharu Funahashi

Kanagawa University

【要約】 金融実務において、派生証券（デリバティブ）の価格付けには（１）モンテカルロ法（MC）や有限差分法（FDM）などの数値計算法、（２）特異摂動法や漸近展開法などの近似手法が用いられている。一方で、ここ数年のファイナンスにおける機械学習（Machine learning（ML））の発展は目覚ましく、数多くの手法が提案されており、数値計算法や近似手法に取って変わる勢いである。本稿では、派生証券（デリバティブ）の価値評価に関する問題に対して、Funahashi（2021）で提案された人工回路網（Artificial neural network,（ANN））を活用したアプローチを中心に紹介し、SABR モデルを用いて実務への応用例を示しながら、その有用性について議論する。

【キーワード】 人工回路網、ディープ・ラーニング、機械学習、近似解、オプション、モンテカルロ法、SABR モデル

【Abstract】 In financial practice, (1) numerical methods, such as Monte Carlo simulation (MC) or finite difference method (FDM), and (2) approximation methods, such as asymptotic expansion or singular perturbation techniques, have been widely used. In contrast, recent progress in machine learning (ML) in the field of finance has shown remarkable development. Various methods have been proposed to price derivatives or calibrate financial asset-pricing models using the artificial neural network (ANN), which seems to be taking the place of the numerical and approximation methods. This note provides an ANN application to the SABR model and discusses the prospects for the ANN application in financial problems.

【Keywords】 Artificial neural network, deep learning machine learning, approximation method, option pricing, monte calro method, SABR model

1. はじめに

金融実務において、流動性の高いヨーロッパ・オプションは市場で観測されるデータにモデルのパラメータを校正（キャリブレーション）するのに用いられる。オプション価格の評価において、実務上最も広く用いられているモデルは Black and Scholes (1973) モデル (BS モデル) であるが、BS モデルでは評価対象となるオプションの権利行使価格 K や満期 T が異なっていたとしてもボラティリティ⁽¹⁾ は一定であることを仮定している。しかし、市場で観測されるインプライド・ボラティリティ (IV) には、スキュー⁽²⁾ やスマイル⁽³⁾ と呼ばれる特性が観測され、 T や K に応じて IV の値が変化するのが一般的である。この事実は、BS モデルでは十分に現実を説明できていないことを示唆している。

このため、実務ではスキューやスマイルを表現可能な局所ボラティリティ・モデル (Local volatility models (LVMs))、確率ボラティリティ・モデル (Stochastic volatility models (SVMs))、或いはそれらのモデルを複合したモデルが好んで使われている。LVMs では Cox and Ross (1976)、Dupire (1997)、Marris (1999) を、SVMs では Hull and White (1987)、Heston (1993)、Hagan et al. (2002)、複合モデルでは Funahashi (2014) を参考文献として挙げておく。

LVMs や SVMs は限られた場合を除き、ヨーロッパ・オプション価格の解析解が存在しない。このことは、原資産価格をモデル化しても、パラメータを市場に校正することができないことを意味しており、モデル選択の実質的なボトルネックとなってきたが、近年では近似手法の開発や計算機能力の向上により、原資産価格が複雑な確率過程に従う場合でもヨーロッパ・オプション価格を算出することができるようになってきている。このあたりの議論は、舟橋 (2015) が詳しい。

近似手法では、漸近展開法が実務で使われることが多い。Hagen et al. (2002) や Fouque et al. (2003) では、特異摂動法を用いて原資産価格が SVMs に従うと過程した下でのデリバティブの価格を算出している。特に、前者が考案するモデルは SABR モデルと呼ばれ、市場で観測される為替やスワップションの IV サーフェイスを構築し補間や補外に使われたり、コンスタント・マチュリティ・スワップ (CMS) の価格付けに用いられており、実務上重要な役割をはたしている。Takahashi (1999) や Takahashi and Takehara (2007) では、小分散理論と呼ばれる漸近理論を発展させ、LVMs や SVMs におけるヨーロッパ・オプション価格の効率的な算出式を導いている。これらの近似手法は、Malliavin 解析における渡辺・吉田理論 (Yoshida (1992)、Watanabe (1987)) と呼ばれる分析に基づき正当化される。Funahashi and Kijima (2015) や Funahashi (2014) では、Wiener-Itô chaos (WIC) 展開を応用した近似手法が提案され LVMs や SVMs におけるヨーロッパ・オプション価格の近似式が与えられている。

しかしながら、これらの手法にも限界がある。近似手法は派生証券価格やそのセンシティブティの効率的な算出を可能にする一方で、近似の次元が上がれば、近似式の算出が煩雑になることに

-
- (1) 原資産価格の収益率の変動を表す指標。
 - (2) 異なる行使価格 K について IV をプロットすると、原資産価格よりも権利行使価格 K が低くなるにつれて上昇する現象をボラティリティ・スキューと呼ぶ。
 - (3) 異なる行使価格 K について IV をプロットすると、下に凸の関数が得られる現象をボラティリティ・スマイルと呼ぶ。

加え、計算負荷も指数的に増加することから、一般には有限次元に限られる。モンテカルロ法を用いたキャリブレーションは、計算負荷が高いことから、利用できるのは計算機に余裕のある一部の銀行や証券会社に限られるし、モンテカルロ法における推定誤差により、市況によっては正しい値が推定できないことも多い。

一方で、ここ数年のファイナンスにおける機械学習 (Machine learning (ML)) の発展は目覚ましく、数多くの手法がモデルの校正や派生証券価格の算出において提案されている。Horvath et al. (2019) では、ANN を駆使した SVMs を市場へ校正する方法が提案されている。本手法の特記すべき点は、校正の過程を (1)Deep approximation と呼ばれるニューラル・ネットワークの学習を行う過程と (2)Calibration と呼ばれる学習の結果を用いてモデルを校正する 2つの過程に分割したことである。こうすることで、直接市場データに SVMs を校正するのに比べて精度や安定性が増すことに加え、時間のかかる Deep approximation は *off-line* で行い、実務上頻繁に行う Calibration とは切り離すことで、*on-line* での calibration を高速に行うことができる。さらに、Horvath et al. (2019) ではボラティリティの従う確率過程を非整数ブラウン運動 (フラクショナル・ボラティリティ・モデル) へ拡張した場合も議論されている。Liu et al. (2019) は、機械学習を用いて Heston (1993) モデルや Bates(1996) モデルのパラメータの校正を行い、Differential Evolution と呼ばれる最適化アルゴリズムを用いると、校正結果の局所的最小解への収束を防ぐことを示した。Hernandez (2017) は、人工回路網を用いてオプション価格を近似することで、モデルのパラメータを高速に市場データに校正することが可能であること証明した。さらに、同論文の中で Inverse mapping と呼ばれる手法も提案している。Inverse mapping では、最初にモデル・パラメータを入力として ANN にヨーロッパ・オプション価格を学ばせた後、逆に (学習結果を用いて生成した) ヨーロッパ・オプション価格を用いて ANN にパラメータを学ばせることで、ANN を用いて市場で観測されるヨーロッパ・オプションの価格から直接パラメータを推定することが可能となる。本手法は、非常に魅力的な手法で将来の実用化が期待されているが、Horvath et al. (2019) で指摘されている通り、(オプション価値を IV に変換する) 逆関数の安定性に問題があることから、現在のところ実務では使われていないようである。この点について、Itkin (2019) は BS モデルを用いて、勾配爆発の問題⁽⁴⁾によりヨーロッパ・オプション価格の ANN での推定が不安定になることを示している。具体的には、オプション価格 C が入力パラメータで、ボラティリティ σ を出力パラメータとしたとき、その勾配は $d\sigma/dC$ で与えられる。ところで、ヨーロッパ・コール・オプションのベガ ($dC/d\sigma$) は権利行使価格 K に対して釣り鐘型をしていることが知られているが、勾配はベガの逆数であることから、権利行使価格が非常に大きい小さいとき大きな値をとり、アット・ザ・マネー (ATM)⁽⁵⁾ において小さな値をとる。このように、勾配の値が急激に変わることが、ANN において勾配爆発や勾配消失として知られる不安定な挙動を誘因し、ANN の学習を阻害する一因となっている。

ANN では、学習を行う *off-line* と学習結果を用いて派生証券を価格付けする *on-line* の過程に分割できて、後者は計算負荷が少なく高速計算が可能であることは既に述べた。一方で、*off-line*

(4) 深層学習 (Deep learning(DL)) において、勾配が前方の層で非常に大きな値となり学習が失敗する現象。

(5) オプションの原資産価格と権利行使価格が等しい状態。

の計算には、学習データを生成するために MC や FDM が用いられるため、計算負荷が高い。Spiegeleer et al. (2018) は ANN の学習を行い、アメリカン・オプションやバリア・オプションの価格を近似したが、アメリカン・オプションでは 100,000、バリア・オプションでは 10,000 の学習データを用いている。しかし、派生証券の価格付けでは学習データ 1 つを作るためには、一般的には試行回数 10,000–1,000,000 回の MC シミュレーションが必要になるし、バリア・オプションなどでは時間の刻み幅 (Time steps) も細かく刻む必要があり計算負荷が高い。さらに、*off-line* での学習はモデルごとに必要になるため、数多くのモデルを使い分ける金融機関では (*off-line* の学習を行うのは年に数回とはいえ) 大きな負荷がかかる。

これらの問題を解決するため、Funahashi (2021) では、これまでの一般的な ANN の手法とは異なり、デリバティブの価格 (C) を直接人工回路網で学習・推定 (C^{ANN}) するのではなく、デリバティブの価格とその WIC 近似より算出した近似解 (\bar{C}) との差 ($D = C - \bar{C}$) を学習させ、その推定結果 (D^{ANN}) を用いてデリバティブの価格を $\bar{C} + D^{ANN}$ として計算した。この計算手法の利点は、デリバティブの価格として $\bar{C} + D^{ANN}$ を用いることで、近似解と ANN の利点である高速計算を保ったうえ、 C 近傍の値 \bar{C} から推定を始めることができるので、 C^{ANN} に比べて安定性が増し、学習の効率が上がる。さらに、近似値 \bar{C} を適切に選ぶことで、推定パラメータに対する D の分散は、 C に比べて圧倒的に小さくなることから、学習データの数や ANN におけるハイパー・パラメータ (ANN における隠れ層やノード) を減らすことができるため、*off-line* の計算負荷を圧縮することができる。さらに、 D を用いることで、学習効率を劣化させていた釣り鐘型のリスク指標が解消される。実際に、Funahashi (2021) では \bar{C} として WIC 近似を用いることでヨーロピアン・オプションやバリアオプションを計算し、従来のモンテカルロ法や人工回路網に比べて高速でかつ安定的に計算できることを数値実験を通して証明しており、学習データ数を 10~100 分の 1 程度に抑えながら、ANN の隠れ層やノードも従来の半数以下に設定しても、派生証券の価格を安定的に計算できることを確認している。

これまで、派生証券の価格評価における人工知能 (AI) の位置付けは、数値計算法或いは近似手法の代替手法であると捉えられてきた。ANN を用いれば、数値計算法とは比較にならないほど高速に派生証券の価格を計算することが可能であるし、モデルや商品ごとに異なった近似手法を開発する必要もない。しかし、金融機関で扱われる派生証券は取引額が大きく、僅かな推定誤差でも巨額の損失を生じる可能性がある。前述した通り、AI にはまだ多くの克服すべき課題が残っており、現状では AI 単体で金融商品を価格付けする能力はない。このため、当面の間は Funahashi (2021) のように、近似手法と AI の技術を組み合わせることで、(先に述べた) 互いの弱点を補完しあい、安定的な運用が模索されると考えている。本稿では、Funahashi (2021) で提案された人工回路網 (Artificial neural network、ANN) を活用したアプローチを中心に紹介し、実務への応用例を示しながら今後の展望について議論する。

本稿の構成は以下の通りである。次節では、本稿で用いるモデルを定義し、表記の整理を行う。3 節では、ANN の簡単な導入を行い、Funahashi (2021) で提案されたアプローチを紹介する。4 節では、3 節で導入した ANN をスワップションの価格付けに応用し、数値例を用いて実務への応用を議論する。最後に、5 節で全体のまとめを行う。

2. モデル

本稿では、原資産価格 S_t とそのボラティリティ v_t は確率微分方程式 (SDE)

$$(2.1) \quad \begin{cases} \frac{dS_t}{S_t} = rdt + \sigma(S_t, v_t)dW_t^S, \\ dv_t = (\theta - \kappa v_t)dt + \gamma(v_t)dW_t^v, \end{cases}$$

に従うと仮定する。ただし、 r は無リスク金利、 θ は長期ボラティリティ (long run volatility)、 κ は平均回帰速度 (speed of mean reversion) を表す定数とし、 $\sigma(s, v)$ は原資産価格 s とそのボラティリティ v の確定値関数、 $\gamma(v)$ は v の確定値関数である。 W^S と W^v はリスク中立測度 \mathbb{Q} の下でのブラウン運動で相関 $dW_t^S dW_t^v = \rho dt$ を持つものとする。

このモデルは、LVMs と SVMs を一般化した汎用的なモデルであり、Funahashi (2014) の中で用いられている。例えば、(2.1) は

- (1) α を正の定数とし、 $\sigma(S_t, v_t) = \alpha$ のときには BS モデル、
- (2) $\sigma(s, v)$ が v に依存しない ($\sigma(s, v) = \sigma(s)$) (2.1) 場合は LVM、
- (3) ϵ を正の定数として $\sigma(s, v) = \sqrt{v}$ 、 $\gamma(v) = \epsilon\sqrt{v}$ と与えたときには Heston (1993) モデルと呼ばれる SVM になる。

ヨーロピアン・コール・オプションとは、満期時点 T における株価 S_T が権利行使価格 K を上回った場合、オプションの購入者に K で株式を購入する権利を与える金融商品である。オプションの購入者は満期 T において S_T が K を上回っていれば、権利を行使することで $S_T - K$ の利得を得るし、逆に満期 T において S_T が K を下回っていれば、権利を放棄すればよいので、ペイオフは 0 となる。これより、ヨーロピアン・コール・オプションのプレミアムは $(x)^+ = \max(x, 0)$ を用いると

$$(2.2) \quad C(T, K) = e^{-rT} \mathbb{E}[(S_T - K)^+]$$

で与えられる。したがって、 $C(T, K)$ を解析的に計算するためには、満期における原資産価格 S_T の分布が必要になる。

例えば、BS モデルでは原資産価格が対数正規分布に従うことから (2.2) は解析的に計算することができる。

$$(2.3) \quad C(T, K) = e^{-rT} \text{Bl}(F_0, r, T, K, \alpha).$$

ただし、 $F_t = S_0 e^{r(T-t)}$ は原資産価格のフォワード価格、 Φ は正規分布の累積密度関数 (CDF) として、Bl はブラック式

$$\text{Bl}(F, r, T, K, \alpha) = F\Phi(d_1(T)) - K\Phi(d_2(T)), \quad d_{1,2} = \frac{\ln\left(\frac{F}{K} \pm \alpha^2 T\right)}{\alpha\sqrt{T}}$$

である。このモデルでは $\sigma(S_t, v_t) = \alpha$ を仮定していることから、 $C(T, K)$ は満期 T や権利行使価格 K が異なっていたとしてもボラティリティは α で一定である。しかし、市場で観測されるインプライド・ボラティリティ (IV) には、スキューやスマイルと呼ばれる特性が観測されており、

T や K に応じて IV の値が変化することが知られている。そこで、実務では現実に近い LVM と SVM の複合モデルが使われている。しかし、一般的には (2.1) で仮定した原資産価格の分布は解析的には求まらない。このため、金融実務では数値計算法や近似手法が用いられている。

SABR モデル (Hagan et al. (2002)) では、 $\theta = \kappa = 0$ 及び

$$(2.4) \quad \begin{cases} \sigma(s, \alpha) = \alpha s^{\beta-1}, \\ \gamma(v) = \nu, \end{cases}$$

を仮定している。このモデルは、市場で観測される為替やスワップションの IV サーフェイスを構築し補完や補外に使われたり、コンスタント・マチュリティ・スワップ (CMS) の価格付けに用いられており、実務上重要なモデルである。SABR モデルでも、(2.2) を解析的に計算することはできないが、特異摂動法を用いた $C(T, K)$ の近似式が広く用いられている。

$$(2.5) \quad C(T, K) \approx e^{-rT} \text{Bl}(F_0, r, T, K, \sigma_{IV}(T, F_0, K)).$$

ただし、

$$\sigma_{IV}(T, f, K) = \begin{cases} \frac{\alpha}{(fK)^{(1-\beta)/2} \left\{ 1 + \frac{(1-\beta)^2}{24} \log^2(f/K) + \frac{(1-\beta)^4}{1920} \log^4(f/K) + \dots \right\}} \left(\frac{z}{x(z)} \right) \xi(T, f, K), & f \neq K \text{ のとき} \\ \frac{\alpha}{f^{(1-\beta)}} \xi(T, f, f). & f = K \text{ のとき} \end{cases}$$

ここで、 $z = \frac{\nu}{\alpha} (fK)^{(1-\beta)/2} \log(f/K)$ 、

$$x(z) = \log \left\{ \frac{\sqrt{1 - 2\rho z + z^2} + z - \rho}{1 - \rho} \right\}$$

$$\xi(T, f, K) = 1 + \left[\frac{(1-\beta)^2}{24} \frac{\alpha^2}{(fK)^{1-\beta}} + \frac{1}{4} \frac{\rho\beta\nu\alpha}{(fK)^{(1-\beta)/2}} + \frac{2-3\rho^2}{24} \nu^2 \right] T + \dots$$

SABR モデルでは、BS モデルと違い、 $\sigma_{IV}(T, f, K)$ が T, K に依存した関数となっていることから、市場で観測されるボラティリティ・スキューやスマイルを復元することができる。

3. 人工回路網を用いた派生証券の価格付け

ここでは、ANN について簡単な導入⁽⁶⁾を行った後、Funahshi (2021) で提案されたアプローチを中心に紹介する。

3-1 人工回路網

人間の脳は、ニューロンと呼ばれる脳を構成する神経細胞が無数に存在し、互いに結合されて集団化することでニューラル・ネットワークが形成されている。ANN は、このメカニズムをヒントに発展してきた数学モデルで、図 1 の左パネルに示す最もシンプルなネットワークを (単純)

(6) ANN の詳細については、例えば岡谷 (2015) を参照。

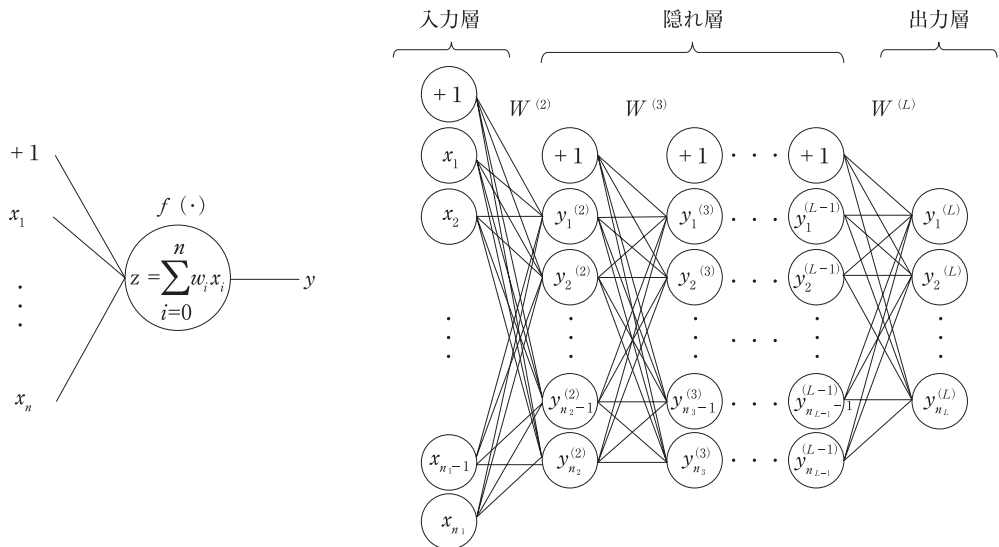


図1 パーセプトロン（左パネル）とL層からなるニューラル・ネットワーク（右パネル）。

パーセプトロンと呼ぶ。パーセプトロンでは、複数の入力信号 x_i ($i = 1, \dots, n$) に重み w_i と呼ばれるパラメータを掛けて足し合わせた値 $z = \sum_{i=0}^n w_i x_i$ をニューロンへ入力し、活性化関数 $f(\cdot)$ と呼ばれる非線形関数を適用して、出力 $y = f(z)$ を得る。なお、図中の丸はノード、ノードを繋ぐ線はエッジと呼ばれている。

パーセプトロンが繋がり、作用し合うことでニューラル・ネットワークが構築される。ニューラル・ネットワークは、入力層、出力層、隠れ層から構成され、層と層の間に重みを配置することで結合の強さを表現し、情報の伝わりやすさを調整している。図1の右パネルに、順伝播型（ニューラル）ネットワークと呼ばれる情報が入力側から出力側に一方向にのみ伝播するニューラル・ネットワークを示す。この例では、ネットワークを構成するユニットが L 層に配置され、隣接する層間でのみ結合を持ち、前の層のユニットの出力が後の層の入力となる形で、結合を通じて信号は左から右へ伝播していく。入力層では、入力 x_i ($i = 1, \dots, n_1$) をそのまま出力する ($y_i^{(1)} = x_i$) と仮定している。 l 層の出力 $\mathbf{y}^{(l)} = f(\mathbf{z}^{(l)})$ は、 $l - 1$ 層から受け取った入力信号 $\mathbf{y}^{(l-1)}$ を重り付けされた合計 $\mathbf{z}^{(l)} = W^{(l)} \mathbf{y}^{(l-1)}$ に変換し、活性化関数 $f(\cdot)^{(7)}$ を作用させることで得られる。ここで、重り W は2階のテンソルとなっていて、 $l - 1$ 層の出力に作用する重りは行列 $W^{(l)} = \{w_{ij}^{(l)}\}_{0 \leq i \leq n_{l-1}, 0 \leq j \leq n_l}$ で表記される。特に、最初の重り w_0 をバイアスと呼び、 $y_0^{(l)} = +1$ に設定されることが多い。

以上の議論から、入力を \mathbf{x} とすると、最終的な出力 $\mathbf{y}(\mathbf{x})$ は

$$(3.1) \quad \mathbf{y}(\mathbf{x}; \mathbf{w}) = f(W^{(L)} f(W^{(L-1)} f(\dots f(W^{(2)} \mathbf{x}))))$$

という活性化関数の深い入れ子状の関数になる。ただし、 \mathbf{w} は W の全要素を並べたベクトルである。順伝播型ネットワークの学習とは、与えられた訓練データ

(7) 活性化関数には $\text{ReLU}(f(x) = \max(x, 0))$ や $\text{SoftMax}(f(x) = e^x / \sum_{j=1}^N e^{x_j})$ など非線形関数が用いられるのが一般的である。本稿では、ReLU を用いる。

$$D = \{(\mathbf{x}_1, \mathbf{y}_1), (\mathbf{x}_2, \mathbf{y}_2), \dots, (\mathbf{x}_N, \mathbf{y}_N)\}$$

をもとに計算される誤差関数 $E(\mathbf{w})$ を最小にするように、ネットワークのパラメータ \mathbf{w} を決定することである。本稿では、誤差関数として

$$E(\mathbf{w}) = \sum_{n=1}^N E_n(\mathbf{w}), \quad E_n(\mathbf{w}) = \frac{1}{2} \|\mathbf{y}_n - \mathbf{y}(\mathbf{x}_n; \mathbf{w})\|^2$$

を用いるものとする。

一般的に、 D はトレーニング・データ（訓練用データ）とテスト・データ（検証用データ）の2つに分割して予測モデルの作成、評価に使われる。テスト・データ数は、全体の15% - 20%に設定されることが多い。

学習の目標は、 $E(\mathbf{w})$ の最小値を与える $\mathbf{w}^* = \arg \min_{\mathbf{w}} E(\mathbf{w})$ を求めることであるが、 $E(\mathbf{w})$ の形状は一般には凸関数ではないので、大域的な最小解は得られない。そこで、極小解を求めるため勾配降下法を用いて W を繰り返し更新する反復計算を行うものとする。すなわち、現在の \mathbf{w} を $\mathbf{w}^{(t)}$ 、負の勾配方向に動かした後の \mathbf{w} を $\mathbf{w}^{(t+1)}$ とすると、

$$\mathbf{w}^{(t+1)} = \mathbf{w}^{(t)} - \epsilon \nabla E(\mathbf{w}).$$

ただし、 $\nabla E(\mathbf{w}) = [\frac{\partial E}{\partial w_1}, \frac{\partial E}{\partial w_2}, \dots, \frac{\partial E}{\partial w_M}]^T$ であり、 w_k は \mathbf{w} の k 成分とする。 $\epsilon > 0$ は、学習係数と呼ばれ \mathbf{w} の更新の大きさを調整する。

勾配降下法を実行するためには、 $\nabla E(\mathbf{w})$ の各要素 $\frac{\partial E}{\partial w_{ji}^{(l)}}$ を効率的に算出する必要がある。しかし、(4.1) から分かる通り、 $w_{ji}^{(l)}$ は $\mathbf{y}(\mathbf{x}; \mathbf{w})$ の中では活性化関数の深い入れ子の中に現れるため、特に入力に近い深い層のパラメータほど、この計算は煩雑になる。

この微分は次の関係式を用いることで、効率よく計算できることが知られている。連鎖規則を用いると

$$(3.2) \quad \frac{\partial E_n}{\partial w_{ji}^{(l)}} = \frac{\partial E_n}{\partial z_j^{(l)}} \frac{\partial z_j^{(l)}}{\partial w_{ji}^{(l)}},$$

となる。(3.2) の右辺最初の値を $\delta_j^{(l)} = \partial E_n / \partial z_j^{(l)}$ (デルタ) と定義すると、

$$(3.3) \quad \delta_j^{(l)} = \sum_k \frac{\partial E_n}{\partial z_k^{(l+1)}} \frac{\partial z_k^{(l+1)}}{\partial z_j^{(l)}} = \sum_k \delta_k^{(l+1)} w_{kj}^{(l+1)} f'(z_j^{(l)}).$$

と変形することができる。ところで、(3.2) の右辺二番目の値は $\partial z_j^{(l)} / \partial w_{ji}^{(l)} = y_i^{(l-1)}$ であるから(3.2) は

$$(3.4) \quad \frac{\partial E_n}{\partial w_{ji}^{(l)}} = \delta_j^{(l)} y_i^{(l-1)}$$

となる。(3.2) は、 $\delta_j^{(l)}$ が $\delta_k^{(l+1)}$ ($K = 1, 2, \dots$) から計算できることを意味しており、最初に出力層の各ユニットのデルタを求めれば、逐次的に任意の層のデルタが計算できる。この方法は、出力層から入力層へとデルタを計算すること（順伝播とは逆）になるので、誤差逆伝搬法と呼ばれ

ている。

3-2 人口回路網の派生証券評価への応用

Funahashi (2021) では、これまでの一般的な ANN の手法とは異なり、デリバティブの価格 (C) を直接人工回路網で学習・推定 (C^{ANN}) するのではなく、デリバティブの価格とその WIC 近似より算出した近似解 (\bar{C}) との差 ($D = C - \bar{C}$) を学習させ、その推定結果 (D^{ANN}) を用いてデリバティブの価格を $CD^{ANN} = \bar{C} + D^{ANN}$ として計算した。この計算手法の利点は、

- (1) *on-line* で必要となる、 D^{ANN} と \bar{C} は計算負荷が小さく高速に計算できる。
- (2) デリバティブの価格として $\bar{C} + D^{ANN}$ を用いることで、 C 近傍の値 \bar{C} から推定を始めることができるので、 C^{ANN} に比べて安定性が増し、学習の効率が上がる。
- (3) 近似値 \bar{C} を適切に選ぶことで、推定パラメータに対する D の分散は、 C に比べて圧倒的に小さくなることから、学習データの数や ANN におけるハイパー・パラメータ (ANN における隠れ層やノード) を減らすことができるため、*off-line* の計算負荷を圧縮することができる。
- (4) D を用いることで、学習効率を劣化させていた釣り鐘型のリスク指標が解消される。

実際に、Funahashi (2021) では \bar{C} として WIC 近似を用いることでヨーロピアン・オプションやバリアオプションを計算し、従来のモンテカルロ法や人工回路網に比べて高速でかつ安定的に計算できることを数値実験を通して証明しており、学習データ数を 10~100 分の 1 程度に抑えながら、ANN の隠れ層やノードも従来の半数以下に設定しても、派生証券の価格を安定的に計算できることを確認している。これまで、モンテカルロ法に代表される数値計算法を用いることで、膨大な計算コスト (大量のサーバーの購入やその維持管理に必要な人件費を含む) を支払っていた金融実務の負荷を大幅に削減できるため、その意義は大きい。なお、WIC 近似を用いることで、 D は多項式に標準正規分布正規分布の累積密度関数 (CDF) や密度関数 (PDF) が乗じられた関数となり、滑らで 2 回微分可能な関数となることが示されている。この結果は、ANN による学習を効率的に行えるだけでなく、Hornik et al. (1990) の定理からグリークス (感応度) の計算も ANN により効率的に学習可能であることを示している。

4. 数値計算

Funahashi (2021) の手法をスワップションの価格付けに応用する。スワップションとは、予め定められた将来の時点 T_α (オプション満期) において、金利スワップ⁽⁸⁾を開始する権利を付与するオプションである。一般的には、オプション満期は最初の変動金利の値を決定する更新日 (reset date) に設定され、その後定期的 (3 ヶ月、半年、1 年ごとが多い) に金利の値洗いを行う。金利の値洗いを行う日をフィクシング・デート $\mathcal{T}^f = \{T_\alpha, T_{\alpha+1}, \dots, T_{\beta-1}\}$ 、支払いが行われる日をペイメント・デート $\mathcal{T}^p = \{T_{\alpha+1}, T_{\alpha+2}, \dots, T_\beta\}$ と区別する。行使価格 K の (ペイヤーズ) スワップションの評価式は、

$$(4.1) \quad \text{PS}(T, K) = A_{\alpha, \beta}(0) \mathbb{E}^{\mathbb{Q}^{\alpha, \beta}} [(S_{\alpha, \beta}(T) - K)^+]$$

(8) もっとも典型的な金利スワップでは変動金利と固定金利の交換を行う。

で与えられることが知られている。ここで、 $A_{\alpha,\beta}(0) = \sum_{i=\alpha+1}^{\beta} \tau_i P(0, T_i)$ はアニュイティと呼ばれ、割引債価格 $P(0, T)$ と金利の適用期間 (year fraction) $\tau_i = T_i - T_{i-1}$ より計算される。 $\mathbb{E}^{\mathbb{Q}^{\alpha,\beta}}$ はアニュイティを基準材とした確率測度のもとでの期待値であり、 $S_{\alpha,\beta}(t) = (P(0, T_\beta) - P(0, T_\alpha)) / A_{\alpha,\beta}(0)$ は時刻 t におけるフォワード・スワップ・レートである。

実務では、 $S_{\alpha,\beta}(t)$ が SABR モデル (2.4) に従うとして、スワップションの価格は計算されるが、(4.1) が $S_{\alpha,\beta}(t)$ の上に書かれたヨーロッパン・オプションであることに注意すれば、この式は近似解 (2.5) を用いて解析的に計算できる。

ANN の推定結果と比較するためのベンチマークとして、数値計算には Schroder (1989) の full truncation Euler scheme を用いるものとし、MC の試行回数は 1,000,000 回、タイム・ステップは 1 年を 250 に分割している。以後、 MC はベンチマーク・プライス、 C^A は Hagan の近似解 (2.5)、 CS^{ANN} は ANN を用いて直接スワップション価格を推定した結果、 CD^{ANN} は \bar{C} に近似解 (2.5) を用いて、Funahashi (2021) の方法で推定したスワップション価格を表すものとする。

4.1 ANN による推定結果

図 2 と図 3 は、それぞれオプション満期 1 年 (スワップ期間 5 年) と 5 年 (スワップ期間 10 年) のスワップション価格を CS^{ANN} 、 CD^{ANN} 、 C^A で計算し、ベンチマーク・プライス (MC) との誤差をヒストグラムで表したものである。トレーニングの学習データの数は 3200、テスト・データは 800 とした。隠れ層は 4 層で、各層のノード数は 32、活性化関数には ReLU を用いた。入力のパラメータ ($f, \alpha, \beta, \nu, \rho, K$) は、表 1 の範囲内で一様乱数を用いて生成したものである⁽⁹⁾。

表 2 に $CS^{ANN} - MC$ 、 $CD^{ANN} - MC$ 、 $C^A - MC$ の平均と分散及び on-line 学習と off-line 推定 (1 商品当たり) に要した計算時間を載せた。図 2、図 3、表 2 より、 CD^{ANN} が CS^{ANN} よりも推定の精度が高いことが分かる。さらに、 CD^{ANN} はベースとなる近似解 C^A よりも近似精度が高く、Hagan の近似解 (2.5) の近似誤差を ANN が上手く補正していることを示している。これらの結果は、Funahashi (2021) の結果と一致している。表 2 より、off-line 計算には時間がかかるが、off-line の計算は高速で計算負荷が小さいことが分かる。

表 1 一様乱数を用いたトレーニング・データの生成範囲

T	f	alpha	beta	nu	rho	K	σ_{IV}
1Y	max	0.100	0.800	1.000	0.800	0.900	0.100
	min	0.002	0.001	0.500	0.001	-0.900	0.002
5Y	max	0.100	0.500	1.000	0.600	0.900	0.100
	min	0.002	0.001	0.500	0.001	-0.900	0.002

表 2 $CS^{ANN} - MC$ と $CD^{ANN} - MC$ の平均と分散及び on-line での学習と off-line での推定 (1 つ当たり) に要した計算時間

	1Y		5Y	
	$CS^{ANN}-MC$	$CD^{ANN}-MC$	$CS^{ANN}-MC$	$CD^{ANN}-MC$
平均	3.40E-04	1.78E-05	1.37E-04	1.49E-06
分散	1.29E-06	2.30E-08	5.27E-06	3.72E-06
off-line 計算時間 (秒)	442	443	467	457
on-line 計算時間 (秒)	1.70E-03	3.71E-03	2.59E-03	1.82E-03

(9) ただし、無意味なパラメータのセットを避けるために、パラメータの範囲として σ_{IV} に制約を加えた。

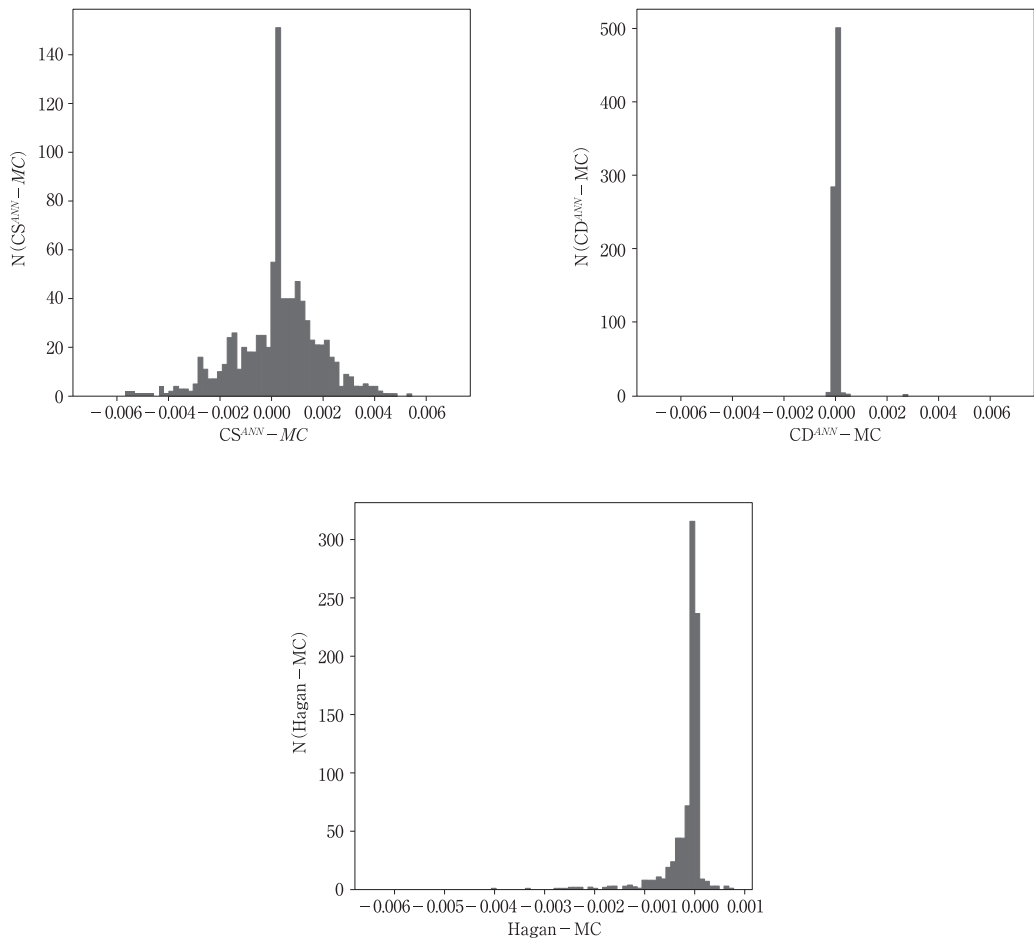


図2 CS^{ANN} 、 CD^{ANN} 、 C^A を用いて計算したスワップション ($T_\alpha = 1Y, T_\beta - T_\alpha = 5Y$) 価格とベンチマーク・プライス (MC) との誤差のヒストグラム。トレーニングの学習データの数 は 3200、テスト・データは 800 とし、入力のパラメータ ($f, \alpha, \beta, \nu, \rho, K$) は、表 1 の範囲内で一様乱数を用いて生成した。

4.2 ANN によるインプライド・ボラティリティの再現性

図 4 の左のパネルは、オプション満期 $T_\alpha = 1$ 、スワップ期間 5 年、 $\tau_i = 0.5$ のスワップションの IV (縦軸) を行使価格 K (横軸) に対してプロットしたものである。 CS^{ANN} と CD^{ANN} は、4.1 において学習したパラメータを用いて計算している。パラメータは、 $f = 0.015, \alpha = 0.373115, \beta = 1, \nu = 0.783108, \rho = -0.24621$ とした。同様に、図 4 の右パネルは、オプション満期 $T_\alpha = 5$ 、スワップ期間 10 年、 $\tau_i = 0.5$ のスワップションの IV (縦軸) を行使価格 K (横軸) に対してプロットしたものである。パラメータは、 $f = 0.02, \alpha = 0.373115, \beta = 1, \nu = 0.783108, \rho = -0.24621$ とした。なお、SABR のモデル・パラメータは恣意性を排除するため、Nkounga (2014) の市場への校正結果 (表 4.7) を用いた。

図 4 より、近似式 (2.5) の精度は行使価格が ATM の周りでは高いが、ATM から離れるほど精度が悪くなっている。また、満期が短い場合 (図 4 の左パネル) は、近似式 (2.5) の精度は高いことが確認できるが、満期が長く (図 4 の右パネル) になれば近似精度が劣化している。一方、

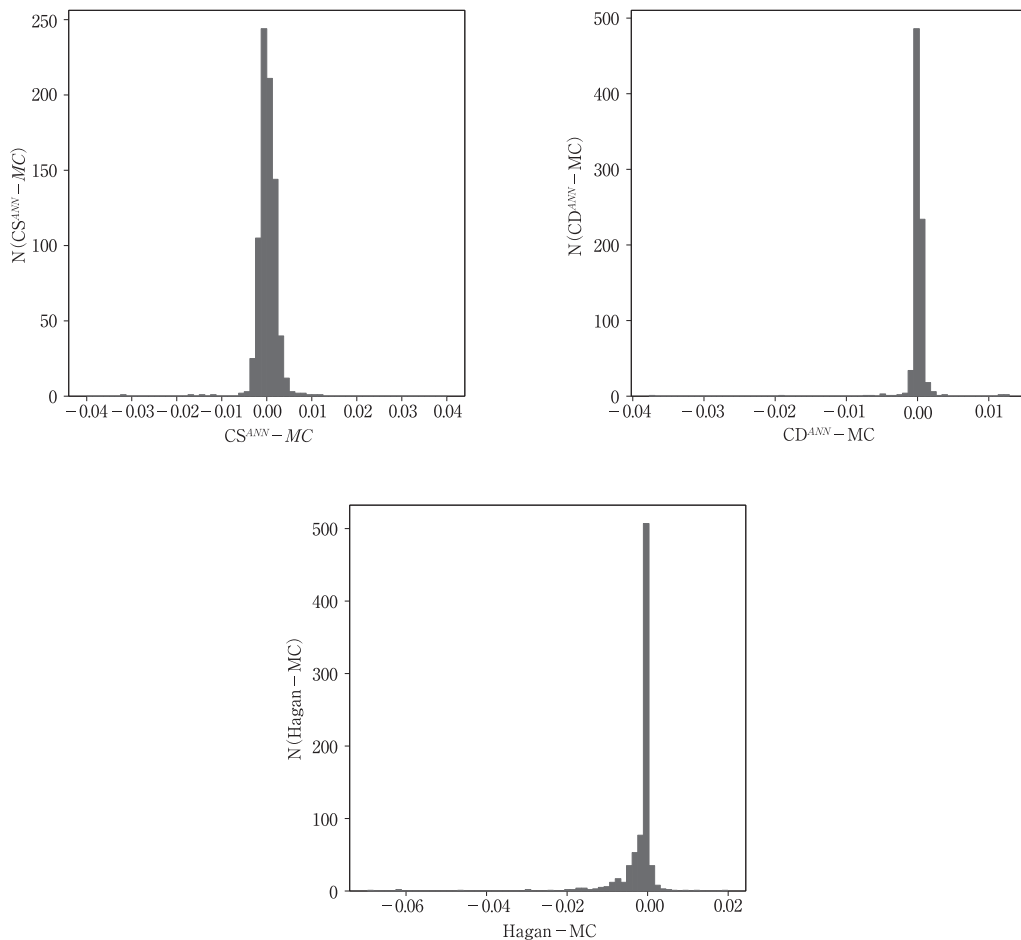


図3 CS^{ANN} 、 CD^{ANN} 、 C^A を用いて計算したスワップション ($T_\alpha = 5Y, T_\beta - T_\alpha = 10Y$) 価格とベンチマーク・プライス (MC) との誤差のヒストグラム、トレーニングの学習データの数は 3200、テスト・データは 800 とし、入力のパラメータ ($f, \alpha, \beta, \nu, \rho, K$) は、表 1 の範囲内で一様乱数を用いて生成した。

CD^{ANN} は、満期に関わりなく近似式 (2.5) の誤差を ANN による学習により補正しており、ベンチマーク (MC) に近づいていることが確認できる。

C^{ANN} は、ANN による学習データが不十分であるため、実務で使うには近似精度が足りない。なお、トレーニング・データ数を 16,000 に増やし C^{ANN} を推定したが、図 4 で計算した 4,000 トレーニング・データより推定した CD^{ANN} の精度には及ばなかった。

ところで、4.1 において、ベンチマークとなる MC の計算時間は、1 商品当たり平均で 1Y で 21.55 秒、5Y で 95.68 秒であった。これらをトレーニング・データ分計算する必要があるので、合計でそれぞれ 86200 秒と 382720 秒の計算時間を要する⁽¹⁰⁾。したがって、トレーニング・データの数を大幅に減らせる Funahashi (2021) の方法は実務上有用であることが確認できる。

(10) 特に、SABR モデルの数値計算は収束が遅く、 MC の試行回数とタイム・ステップを大きくとる必要があることが知られている。例えば Chen et al. (2012) を参照。

5. 結論

本稿では、Funahashi (2021) の手法を SABR モデルに適用することで、Hagan et al. (2002) の近似式では、近似精度が十分でなかった行使価格が ATM から離れた deep in the money や deep out of the money でも、ANN による補正により高い精度で近似することが可能であることが示された。一方で、従来通りオプション価格を直接 ANN で近似させると、実務で要求される近似精度に収めるためには、大量のトレーニング・データを要し計算負荷が極めて高いことも議論した。

これまで、派生証券の価格評価における人工知能 (AI) の位置付けは、数値計算法或いは近似手法の代替手法であると捉えられてきた。しかし、金融機関で扱われる派生証券は取引額が大きく、僅かな推定誤差でも巨額の損失を生じる可能性がある。ANN にはまだ多くの克服すべき課題が残っており、現状では ANN 単体で金融商品を価格付けするには計算コストが掛かり過ぎる。このため、当面の間は Funahashi (2021) のように、近似手法と ANN の技術を組み合わせることで、互いの弱点を補完しあい、トレーニング・データ数やハイパー・パラメータを抑えながら、安定的な運用が模索されることが考えられる。

今後の研究としては、ANN の技術を用いて、原資産価格が負金利に対応した SABR モデルの近似解の開発や、SABR モデルの wings (deep in the money や deep out of the money) における近似精度の向上が考えられる。さらに、解析的にも数值的にも計算負荷が大きく実務的な運用が遅れている、フラクショナル・ボラティリティ・モデルへの ANN 技術の応用も期待される。

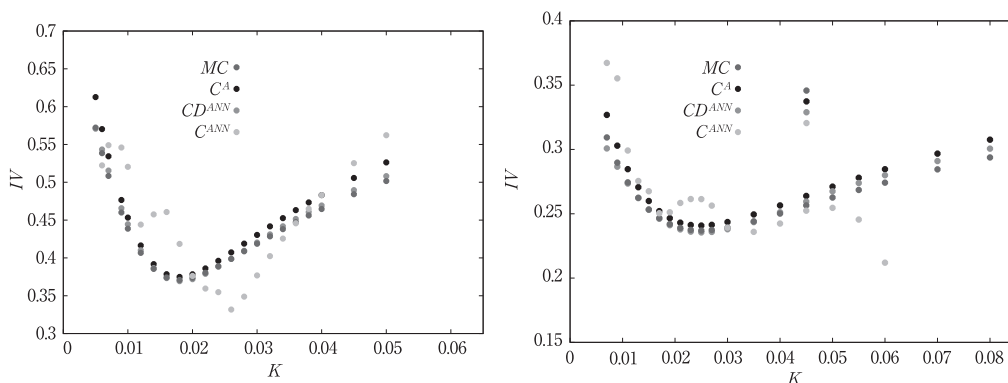


図4 左パネルは、オプション満期 $T_\alpha = 1$ 、スワップ期間 5 年、 $\tau_i = 0.5$ のスワップションの IV。パラメータは、 $f = 0.015, \alpha = 0.373115, \beta = 1, \nu = 0.783108, \rho = -0.24621$ 。右パネルは、オプション満期 $T_\alpha = 5$ 、スワップ期間 10 年、 $\tau_i = 0.5$ のスワップションの IV。パラメータは、 $f = 0.02, \alpha = 0.373115, \beta = 1, \nu = 0.783108, \rho = -0.24621$ 。 CS^{ANN} と CD^{ANN} は、4.1 において学習したパラメータを用いて計算している。

●参考文献

- Bates, D. S. (1996), “Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche mark options,” *The Review of Financial Studies*, **9**(1), 69–107.
- Black, F. and M. Scholes (1973), “The pricing of options and corporate liabilities,” *Journal of*

- Political Economy*, **81**, 637–654.
- Chen, B., W. C. Oosterlee, and H. V. D. Welde (2012), “A low-bias simulation scheme for the SABR stochastic volatility model,” *International Journal of Theoretical and Applied Finance*, **15**(2), 1250016.
- Cox, J. C. and S. A. Ross (1976), “The valuation of options for alternative stochastic processes,” *Journal of Financial Economics*, **3**, 145–166.
- Dupire, B. (1997), “Pricing and Hedging with Smiles,” *Cambridge University Press*, January, 103–111.
- Marris, D. (1999), “Financial Option Pricing and Skewed Volatility,” MPhil thesis, University of Cambridge.
- Fouque, J. P., G. Papanicolaou, K. R. Sircar, and K. Solna (2003), “Singular Perturbations in Option Pricing,” *Journal on Applied Mathematics*, **63**(4), 1648–1681.
- Funahashi, H. (2014), “A chaos expansion approach under hybrid volatility models,” *Quantitative Finance*, **14**, 1923–1936.
- Funahashi, H. and M. Kijima (2015), “A chaos expansion approach for the pricing of contingent claims,” *Journal of Computational Finance*, **18**, 27–58.
- Funahashi, H. (2019), “Artificial neural network for option pricing with and without asymptotic correction,” *Quantitative Finance*, forthcoming.
- Hagan, P., D. Kumar, and A. Lesniewski (2002), “Managing smile,” *Wilmott magazine*, September, 84–108.
- Hernandez, A. (2017), “Model calibration with neural networks,” *Risk*, June, 1–5.
- Heston, S. (1993), “A Closed Form Solution for Options with Stochastic Volatility with Applications to Bond and Currency Options,” *Review of Financial Studies*, **6**, 327–343.
- Horvath, B., A. Muguruz, and M. Tomas, (2019), “Deep learning volatility,” working paper.
- Hornik, K., M. Stinchcombe, H. White, (1990), “Universal Approximation of an Unknown Mapping and Its Derivatives Using Multilayer Feedforward Networks,” *Neural Networks*, **3**, 551–560.
- Hull, J. and A. White (1987), “The Pricing of Options on Assets with Stochastic Volatilities,” *Journal of Finance and Quantitative Analysis*, **3**, 281–300.
- Itkin, A. (2019), “Deep learning calibration of option pricing models: some pitfalls and solutions,” working paper.
- Liu, S., A. Borovkyh, L. Grzelak, C. Oosterlee (2019), “A neural network-based framework for financial model calibration,” working paper.
- Nkouna, M., J. T. Johnattan (2014), Building Interest Rate Curves and SABR Model Calibration, PhD thesis, Stellenbosch University.
- Schroder, M. (1989), “Computing the constant elasticity of variance option,” *Journal of Finance*, **44**, 211–219.
- Spiegeleer, J. D., D. Madan, S. Reyners, W. Schoutens (2018), “Machine learning for quantitative finance: fast derivative pricing, hedging and fitting,” *Quantitative Finance*, **18**(10), 1–9.
- Takahashi, A. (1999), “An Asymptotic Expansion Approach to Pricing Financial Contingent Claims,” *Asia-Pacific Financial Market*, **6**, 115–151.
- Takahashi, A. and K. Takehara (2007), “An Asymptotic Expansion Approach to Currency Options with a Market Model of Interest Rates under Stochastic Volatility Process of Spot Exchange Rates,” *Asia-Pacific Financial Market*, **14**, 69–121.
- Watanabe, S. (1987), “Analysis of Wiener functionals (Malliavin calculus) and its applications to heat kernels,” *Annals of Probability*, **15**, 1–39.
- Yoshida, N. (1992), “Asymptotic expansions of maximum likelihood estimators for small diffusions via the theory of Malliavin-Watanabe,” *Probability Theory and Related Fields*, **92**, 275–311.

岡谷貴之 (2015), 『深層学習』, 講談社。

舟橋秀治 (2015), 「カオス展開法を用いた金融派生証券の価格付け」, 博士論文。