

動的環境における強化学習のステップサイズパラメータ調整法

野田五十樹^{†a)}

Gradient Descent Adaptation of Stepsize Parameter in Reinforcement Learning for Dynamic Environments

Itsuki NODA^{†a)}

Abstract. 動的な環境において、強化学習のステップパラメータを、学習の過程において調整する方法を提案する。一般の強化学習では、学習対象である環境は一定と仮定されており、ステップパラメータは雑音などの効果を除去するために徐々に単調に減衰させることが多い。しかしマルチエージェント技術が適用されるような環境では、環境は絶えず変化する場合が多く、その変化の度合いに応じてステップパラメータを調整する必要がある。本研究では、この調整方法を強化学習の過程で行う方法を提案し、その数学的解析および実験により、調整法の持つ意味や有効性を検討する。

Keywords. 強化学習、ステップサイズ、動的環境、適応

1. まえがき

エージェントの学習方法として広く用いられている強化学習では、エージェントが活動する環境設定あるいは問題設定は変動しないと仮定されていることが多い。すなわち、エージェントは動作することで環境に何らかの変化が生じるが、その変化の生じ方や報酬の与えられ方は変化しないものとして扱われる。例えば、強化学習の1つであるQ学習[3]では、以下のような式によりQ値を更新する。

$$Q_{t+1}(s_t, a_t) = (1 - \alpha)Q_t(s_t, a_t) + \alpha(r_t + \gamma \max_{a'} Q_t(s_{t+1}, a')) \quad (1)$$

この式で使われるステップサイズパラメータ α は0~1の値をとるが、一般に0に漸近させる形で変化させる。これは、状態遷移における確率的揺らぎや報酬に重畳される雑音の成分を省き、Q値を真の期待報酬値に近づけるためである。そして、Q値が十分に真の値に近づいたと思われる時点で学習を止め、実際の運用ではその結果のみを用いて学習なしに行動決定を行う形でエージェントを動作させることが一般的で

ある。ここで重要なのは、この真の期待報酬値が変化しないことが前提になっている点である[1]。

一方、実世界における問題や開放系の問題では、エージェントの環境自体は徐々にあるいは急激に変化することが一般的である。例えば為替取引などの経済活動では、関係する要因が多数にわたるための複雑さに加え、そこに参加して活動している他のエージェントの行動原理自身が時間的に変化するため、同じ状況でも「期待報酬が一定である」という仮定は置きにくい。このような場合、エージェントの学習を「完了」させることはせず、継続的に学習を行わせ環境の変化に適切にエージェントの行動を追従させて行くことが重要になる。つまり、式(1)の α を0に近づけるといった形で学習を収束させるのではなく、絶えず真の値や最適行動を探査できるエージェントの学習方式が望まれる。

このような問題に対し[2]では、強化学習の学習過程の初期の過渡的な部分について、雑音を除去しながら望ましい値に収束させる方法を提案している。また、[4]ではエージェントが獲得する報酬の分散に着目した学習方法を提案している。しかしいずれにしても、学習対象は最終的には不変であり、学習プロセスもいつか「完了」という前提がなされている。

これに対して本稿では、上記のような真の期待報酬が徐々に変化するような動的な環境において学習し

[†] (独) 産業技術総合研究所 情報技術研究部門、〒305-8568 茨城県つくば市梅園 1-1-1

a) E-mail: i.noda@aist.go.jp

続けるエージェントを仮定し、学習で用いられるパラメータをいかに調整して環境の変化に適切に追従できるようにするか、という問題を取り上げる。具体的には、式 (1) のステップパラメータ α に着目し、この値を環境の変化に対応させる方法について提案する。

2. 指数平滑移動平均とステップパラメータ

2.1 指数平滑移動平均

Q 学習や TD 学習などの強化学習では、未知の環境における状態や行動の価値を、実際に行動して得られた報酬を元に推定し、そこで得られた推定価値をもとに行動を決定するかたちで学習が進む。この価値の推定で用いられる式を一般化すると、下記のような指数平滑平均 (Exponential Moving Average, EMA) の式で表される。

$$\tilde{x}_{t+1} = (1 - \alpha)\tilde{x}_t + \alpha x_t \quad (2)$$

ここで、 x_t および \tilde{x}_t は経験によって実際に観測された値 (報酬 r_t など) およびその推定値であり、時刻 t によって更新されていく。また、 α はステップパラメータである。この α は、直近の観測値 x_t をいかに重視するか、あるいは、どの程度長い時間の移動平均として推定値 \tilde{x}_t を求めるかを示している。一般に、式 (2) で求められる \tilde{x}_t は、 x_t について $T = \frac{2}{\alpha} - 1$ の時間幅の単純移動平均を近似していることが知られている。

2.2 ランダムウォークする値への追従

ここで、観測値の系列 $\{x_t\}$ が下記のように、求めるべき真の値の系列 $\{s_t\}$ に雑音が重畳したものと考える。

$$x_t = s_t + \epsilon_t \quad (3)$$

ただし ϵ_t は平均 0、標準偏差 σ_ϵ の乱数とする。

さらに、真の値の系列 $\{s_t\}$ は下記のようにランダムウォークで変化する値とする。

$$s_{t+1} = s_t + v_t$$

ただし、 v_t は平均 0、標準偏差 σ_v の乱数とする。

この場合、以下の補題と定理が成立する。

[補題 1]

式 (3) で与えられ観測値 x_t に対し、式 (2) で求めた推定値 \tilde{x}_t の誤差 $\delta_t = \tilde{x}_t - x_t$ の二乗平均は、以下の式で与えられる。

$$\mathbf{E}(\delta_t^2) = \frac{1}{2 - \alpha} (2\sigma_\epsilon^2 + \frac{1}{\alpha}\sigma_v^2) \quad (4)$$

□

[定理 1]

誤差 $\delta_t = \tilde{x}_t - x_t$ の二乗平均を最小にするステップサイズ α は以下の式で与えられる。

$$\alpha = \frac{-\gamma^2 + \sqrt{\gamma^4 + 4\gamma^2}}{2} \quad (5)$$

ただし、 $\gamma = \frac{\sigma_v}{\sigma_\epsilon}$ である。

□

この定理は、入力される観測値がランダムウォークとそれと独立な雑音から構成されている場合、ランダムウォークによる真の値の変化への追従と移動平均による雑音成分の平滑化をバランスさせるためには、式 (5) によりステップサイズパラメータを定めればよいことを示している。

2.3 再帰的指数平滑移動平均と偏微分

式 (5) では、ランダムウォークや雑音の標準偏差が既知であることを仮定している。しかし一般にこれらの値がわからなかったり、あるいは動的に変化する場合があり得る。そこで、以下では、観察された時系列 $\{x_t\}$ が与えられている場合に、それに最も適した α を適応的に決定する方法を考える。

まず、以下のように式 (2) を再帰的に適用した再帰的指数平滑移動平均 (Recursive Exponential Moving Average, REMA) $\xi_t^{(k)}$ を導入する。

$$\begin{aligned} \xi_t^{(0)} &= x_t \\ \xi_{t+1}^{(1)} &= \tilde{x}_{t+1} = (1 - \alpha)\tilde{x}_t + \alpha x_t \\ \xi_{t+1}^{(k)} &= (1 - \alpha)\xi_t^{(k)} + \alpha \xi_t^{(k-1)} \\ &= \alpha \sum_{\tau=0}^{\infty} (1 - \alpha)^\tau \xi_{t-\tau}^{(k-1)} \end{aligned} \quad (6)$$

この時、以下の補題が成立する。

[補題 2]

REMA $\xi_t^{(k)}$ の α による 1 階偏微分は次式で与えられる。

$$\frac{\partial \xi_t^{(k)}}{\partial \alpha} = \frac{k}{\alpha} (\xi_t^{(k)} - \xi_t^{(k+1)}) \quad (7)$$

□

[定理 2]

EMA \tilde{x}_t ($k = 1$ の REMA) の k 階偏微分は下記で与えられる。

$$\frac{\partial^k \tilde{x}_t}{\partial \alpha^k} = (-\alpha)^{-k} k! (\xi_t^{(k+1)} - \xi_t^{(k)}) \quad (8)$$

□

2.4 再帰的指数平滑移動平均によるステップサイズ勾配降下法

定理 2 により α による \hat{x}_t の導関数が求まるので、誤差 δ_t の二乗誤差を逐次的に極小化する勾配降下法を導くことができる。ただし、定理 2 では高階の導関数も求めることができるので、より精度の高い降下法を求めることができる。よって、この高階導関数を用いた勾配降下法を再帰的指数平滑移動平均によるステップサイズ勾配降下法 (Gradient Descent Adaptation of Stepsize by Recurrent Exponential Moving Average, GDASS-by-REMA あるいは GDASS) と呼ぶ。

今、仮に $\alpha \rightarrow \alpha + \Delta\alpha$ と変化させた場合の \hat{x}_t の変化分を $\Delta\hat{x}_t$ とすると、

$$\begin{aligned}\Delta\hat{x}_t &= \sum_{k=1}^{\infty} \frac{1}{k!} \frac{\partial^k \hat{x}_t}{\partial \alpha^k} \Delta\alpha^k \\ &= \sum_{k=1}^{\infty} (-1)^k \left(\frac{\Delta\alpha}{\alpha}\right)^k (\xi_t^{(k+1)} - \xi_t^{(k)}) \quad (9)\end{aligned}$$

さらにまた、一般の $\xi_t^{(k)}$ の変化についても、補題 2 の結果を用いて 1 次の Taylor 展開を行うと以下のようになる。(注1)

$$\Delta\xi_t^{(k)} \simeq k \left(\frac{\Delta\alpha}{\alpha}\right) (\xi_t^{(k)} - \xi_t^{(k+1)}) \quad (10)$$

これらの展開式には、以下の意味がある。

(1) 高階導関数が REMA により求められることで、 α を比較的大きく動かしても、それによる \hat{x}_t や δ_t の変化分を精度良く近似できる。このため、同じ勾配降下法でも、修正幅を大きくとることができる。

(2) 導関数を求めるのに必要な変数 $\xi_t^{(k)}$ についても、同様に導関数を求めることができるので、 α を修正した際にこれらの変数の値も修正できる。よって、連続的に修正を行っても、精度が保たれる。

もちろん、REMA を無限の k について求めることはできないため、有限の k で打ち切る必要がある。しかし REMA の計算自体は単純であり、必要な精度に応じて k の上限 k_{\max} を定めることで、実質的な勾配降下法を実現できる。

具体的には、以下のような手順で学習を行う。

初期化: $\forall k \in \{0 \dots k_{\max} - 1\} : \xi^{(k)} \leftarrow x_0$

while forever do

観測データを x とする。

for $k = k_{\max} - 1$ **to** 1 **do**

$$\xi^{(k)} \leftarrow (1 - \alpha)\xi^{(k)} + \alpha\xi^{(k-1)}$$

end for

$$\xi^{(0)} \leftarrow x$$

$$\delta \leftarrow \xi^{(1)} - x$$

$\frac{\partial \xi^{(1)}}{\partial \alpha}$ を式 (8) により求める。

δ および $\frac{\partial \xi^{(1)}}{\partial \alpha}$ から、 α の変化分 $\Delta\alpha$ を決定。

for $k = 1$ **to** $k_{\max} - 1$ **do**

式 (9) および式 (10) により $\Delta\xi^{(k)}$ を求める。

$$\xi^{(k)} \leftarrow \xi^{(k)} + \Delta\xi^{(k)}$$

end for

$$\alpha \leftarrow \alpha + \Delta\alpha$$

end while

この中で、 $\Delta\alpha$ を決定する方法にはいくつか考えられる。通常の勾配降下法と同じく、基本的には $\delta \frac{\partial \xi^{(1)}}{\partial \alpha}$ が正ならば $\Delta\alpha < 0$ 、負であれば $\Delta\alpha > 0$ とすればよい。ただし、下記の点に注意する必要がある。

- α は $[0, 1]$ 範囲内の実数である必要がある。
- 式 (8) は $\alpha = 0$ で特異点となるため、 $\alpha 0$ をあまり 0 に近づけてはいけない。

これらを踏まえ、次節の実験では、以下の方法で $\Delta\alpha$ を決めている。

$$\begin{aligned}\gamma'_{\text{old}} &\leftarrow \sqrt{\frac{\alpha^2}{1 - \alpha}} \\ \lambda &\leftarrow -\bar{\lambda} \cdot \text{sign}\left(\delta \frac{\partial \xi^{(1)}}{\partial \alpha}\right) \\ \gamma'_{\text{new}} &\leftarrow \exp(\log(\gamma'_{\text{old}}) + \lambda) \\ \alpha_{\text{new}} &\leftarrow \frac{-\gamma_{\text{new}}'^2 + \sqrt{\gamma_{\text{new}}'^4 + 4\gamma_{\text{new}}'^2}}{2} \\ \Delta\alpha &\leftarrow \alpha_{\text{new}} - \alpha\end{aligned}$$

これは、 α を直接増減させるのではなく、式 (5) の関係に注目し、 α を γ に変換したあと、その γ の対数軸上で一定幅 $\bar{\lambda}$ 増減させるという方法となっている。このため、 α 上での増減の大きさは 0 および 1 付近では細かく、0.5 付近では大きくなっている。

3. 実験

3.1 実験 1: 最適な α の学習

まず、上で述べた α の更新方法により、式 (5) で示した最適な値が得られるかどうかを、実験により確認

(注1) : 付録に示した高階導関数を用いれば、高次の Taylor 展開を求めることもできる。

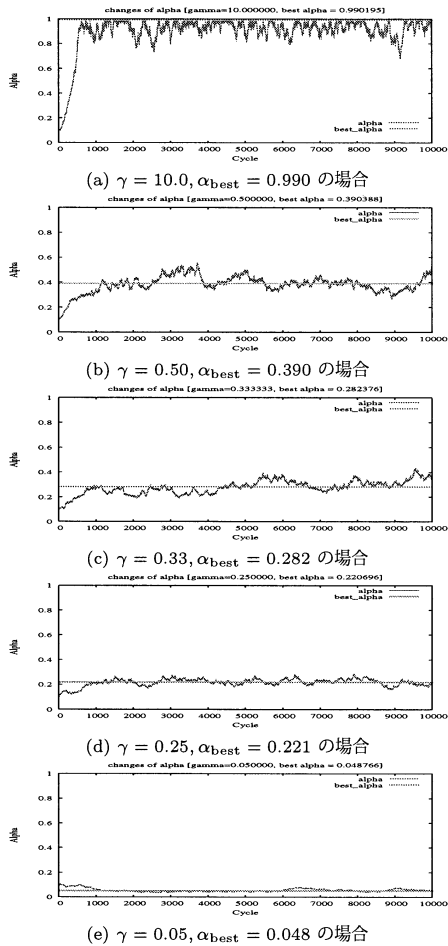


図1 実験1: 標準偏差比 γ の違いによる α の学習過程の変化
 α_{best} および 'best.alpha' は、式 (5) による最適な α を示す。

する。

図1は、異なる γ を持つ観測データ x_t を用いて α を適応させた場合の実験結果である。各パラメータについては、以下のように設定した。

	σ_s	σ_ϵ	γ	α_{best}
(a)	0.01	0.001	10.00	0.990
(b)	0.01	0.020	0.50	0.390
(c)	0.01	0.030	0.33	0.282
(d)	0.01	0.040	0.25	0.221
(e)	0.01	0.200	0.05	0.048

これらのグラフは、学習回数 (横軸) が進むのに従って、 α がどのように変化したかを示している。また、

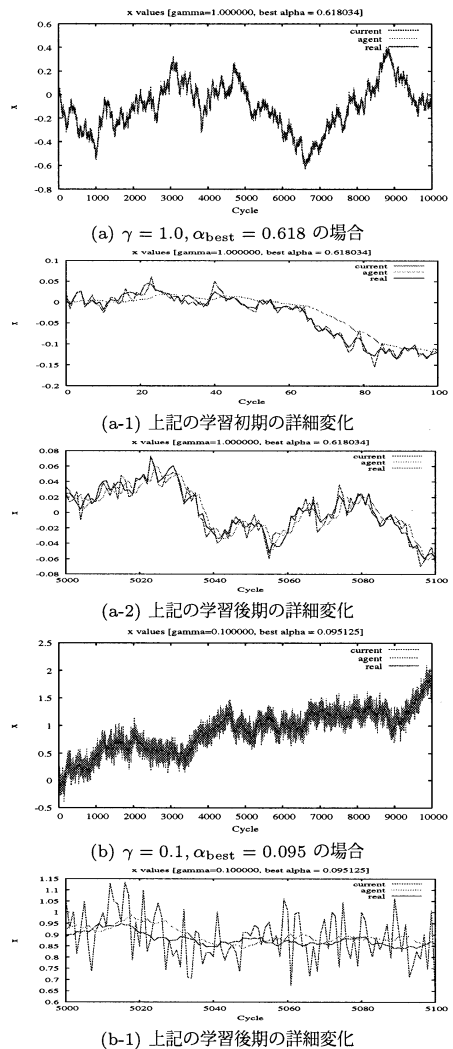


図2 実験1: 予測値 \tilde{x}_t の様子
 図中、'current' は観測値 x_t 、'agent' は EMA による推定値 \tilde{x}_t 、'real' は真の値 s_t

グラフ中に描かれている水平の直線は、式 (5) で求めた最適な α_{best} である。この α の変化を見ると判るように、学習を経るに従い、 γ に応じた最適なステップサイズを獲得できていることが判る。ただし、観測データに重畳されるノイズの影響により、最適値に収束していくのではなく、最適値の周辺で変動を続けるようになる。

また、図2は、学習の過程で得られている入力デー

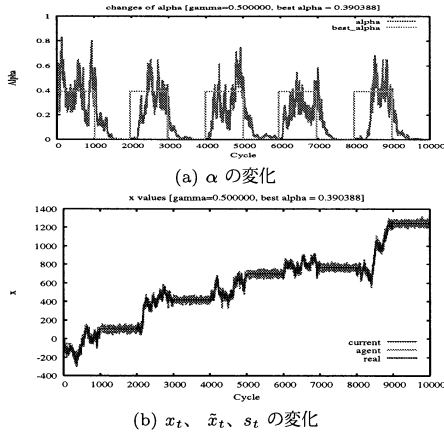


図3 実験 2: γ が矩形波状に変化する場合。
 $\gamma = 0.5$ と $\gamma = 0.0005$ を 1000 ステップごとに繰り返す。

タの推定値 \hat{x}_t の変化である。(a) は真の値のランダムウォークが相対的に強い ($\gamma = 1.0$) 場合、(b) は雑音相対的に大きい ($\gamma = 0.1$) の場合である。また、(a-1)・(a-2) は (a) の学習の初期・後期の詳細を、(b-1) は (b) の学習後期の詳細を示している。まず (a-1) では、学習が進んでおらず α が小さすぎるため、真の信号あるいは観測データに対し、かなり鈍った変化を \hat{x}_t は示している。一方、(a-2) では α が適切な値になったため、 \hat{x}_t が真の信号の変化に追従できていることが判る。一方、雑音成分が多い (b) あるいは (b-1) では、入力データに含まれる雑音成分を平滑化し、真の値 s_t に近い値を \hat{x}_t が示していることがわかる。

以上のように、与えられた観測データに含まれる真の値のランダムウォークの大きさと雑音成分の大きさに応じたステップサイズを、GDASS により獲得できていることがわかる。

3.2 実験 2: γ が矩形波状に変化する場合

次に、最適な α を決める γ が途中で変化する場合にどのように追従できるかを示すため、 γ が矩形波状に時間的に変化する観測データを用いて学習実験を行った。

図 3 は、 γ の値が 0.5 と 0.0005 に 1000 ステップごとに変化する観測データ x_t を用いて、GDASS による α の学習及び EMA \hat{x}_t による推定を行った結果である。この図の (a) は、学習による α を、変化する γ による最適ステップサイズの変化と共に示したものである。図中、矩形に変動しているのが最適ステップサ

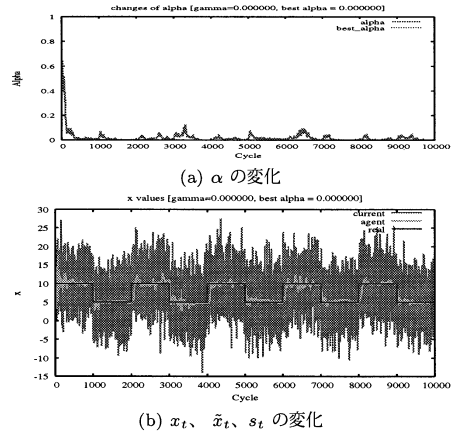


図4 実験 3: 真の値が矩形振動で変化する場合。
 s_t が 0.0 と 5.0 を 1000 ステップごとに繰り返す。
 $\sigma_\epsilon = 5.0$

イズであり、矩形の上値・下値は各々 0.39 と 0.0005 である。この図から、 α は最適ステップサイズを追うように変化していることが判る。また、(b) は同じ実験における \hat{x}_t の変化を観測データ x_t と真の値 s_t と共に示したものである。こちらでは、 γ が小さな時間帯では雑音成分を平滑化して真の値に近い値を \hat{x}_t が示しているのに対し、 γ が大きな時間帯では大きく変化している真の値に追従できていることがわかる。すなわち、学習の最中に環境が変化し γ が変動しても、 α も追従して適切な推定値を出力できるようになっている。

3.3 実験 3: 真の信号が矩形波状に変化する場合

強化学習などで指数平滑移動平均の式が用いられるのは、報酬などの観測データに加わる雑音成分を取り除き、真の値を求めるためである。一方、本稿の目的は環境の変化により真の値が変化した場合、それを検知して追従できる方法を構築することである。GDASS がこの目的を果たしていることを確認するために、真の値が定期的に矩形波形で変化するデータをもちいた実験を行った。

図 4 は、1000 ステップごとに 5.0 と 10.0 に交互に変化する真の値 s_t に標準偏差が 5.0 のガウス雑音を重畳した観測データ x_t を作成し、それを用いて \hat{x}_t および α の学習を行わせた結果である。この図の (b) から判るように、GDASS は大きな雑音成分をきれいに除去し、精度良く s_t の変化に追従できていることが判る。これに対応する α の変化 ((a) 図) を見ると、

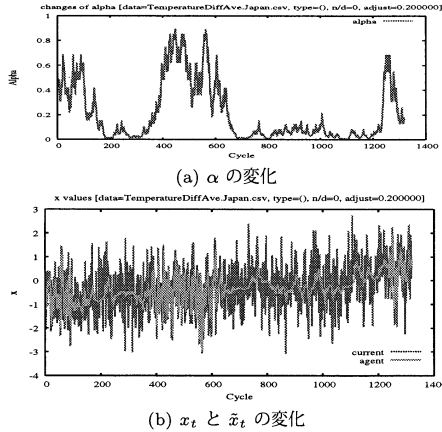


図 5 日本の平均気温を観測値 x_t として用いた学習結果

α はほぼ 0 に張り付いているが、 s_t が変化するタイミング ($t = 1000, 2000, \dots$) に近いところで α が少し盛り上がっていることが判る。指数平滑移動平均における α の意味 (α が大きければ直前の観測に追従し、 α が小さければ長期の時間平均をとる) と照らし合わせると、学習におけるこの α の変化は、 s_t が変化するところでは直前追従に振れ、安定しているとき長時間平均で雑音を除去しようとしていることになる。すなわち、GDASS が環境の変化に応じて雑音除去と変化追従を使い分けて両立していることがわかる。

3.4 実データを用いた例

最後に、実データに本手法を適用した結果を示す。

図 5 は、1898 年から 2007 年にかけての日本の毎月の平均気温 (平年比) のデータを観測データ x_t として、GDASS を適用した時の実験結果である。これを見ると、多くのところで α は小さいが、サイクルが 0 ~ 200、400 ~ 700 および 1300 付近においては大きくなっていることが判る。すなわち、この期間の気温変動は単なる雑音成分ではなく、なんらかの本質的な変化が大きくなっているという可能性を、この学習結果は示している。

同様に、外国為替相場 (円-ドル、1990 年 ~ 2008 年、日足終値) のデータを用いて学習実験を行った結果を図 6 に示す。この図のうち、(a-1)・(a-2) は円-ドルレートそのもの、(b-1)・(b-2) は前日との差を観測データとして学習を行わせたものである。(a-1) から α が 1 に近いところを保ちながら、ところどころ下落していることがわかる。つまり、レートの値そのもの

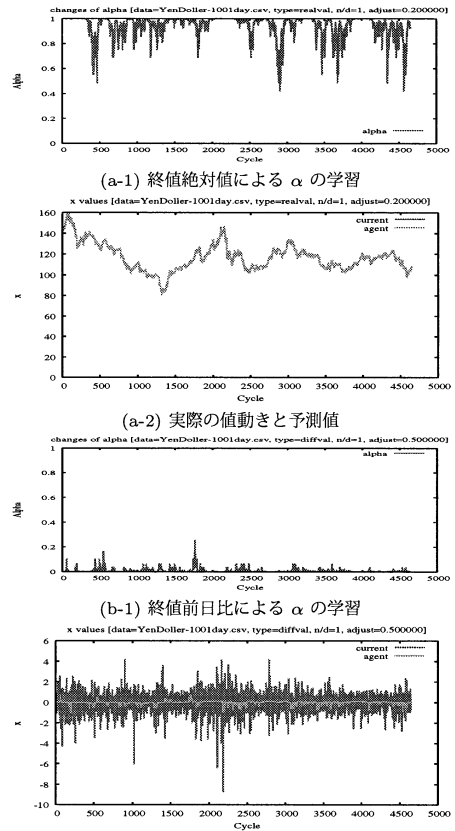


図 6 円-ドル相場を用いた学習結果

はランダムウォークそのものであり、雑音成分は少ないことがわかる。一方、(b-1) では α はほぼ 0 であり、ところどころ盛り上がっている。この変化は図 4 に示した実験 3 の結果に近いとみなせる。すなわち、推定値 \hat{x}_t は、観測データに含まれる大きな雑音成分のなかから、昇り相場・下り相場を検出している可能性がある。

4. まとめ

本稿では、徐々に変化する環境の中で、エージェントが学習を継続しながら環境の変化に適応していくという前提のもと、学習パラメータであるステップサイズ α と環境から得られる観測データの関係を明らかにした。その結果を元に、 α を環境の変化に追従させる方法として再帰的指数平滑移動平均によるステップサイズ勾配降下法 (GDASS) を提案した。この方法の特徴は、通常の勾配降下法と異なり、高階の導関数をシステマティックに求めることができる点である。このため、勾配法による α の変更量を比較的大きく取ることができ、さらに、修正の際に利用する値も同時に修正する方法を提供できる。

GDASS の性能を検証する実験では、GRASS が最適な α を獲得できること、環境の変化に追従して α を調整できることを示した。ただし、3.4 節で示した実データによる実験では、獲得された α の変化が、現実のどのような側面を反映しているのかは、現在のところまだ不明である。今後は、これらの分析などを行っていく必要がある。

謝辞 本研究を進めるに当り貴重なアドバイスを頂きました産業技術総合研究所 和泉潔氏、CMD Lab 松井宏樹氏、尹熙元氏に感謝します。

文 献

- [1] EVEN-DAR, E. and MANSOUR, Y. Learning rates for Q-learning, *Journal of Machine Learning Research*, 5 (Dec. 2003), 2003.
- [2] GEORGE, A. P. and POWELL, W. B. Adaptive stepsizes for recursive estimation with applications in approximate dynamic programming, *Machine learning*, 65, 1 (2006), 167–198.
- [3] SUTTON, R. S. and BARTO, A. G. *Reinforcement Learning: An Introduction*, MIT Press, Cambridge, MA (1998).
- [4] 佐藤誠, 木村元, 小林重信 報酬の分散を推定する TD アルゴリズムと Mean Variance 強化学習法の提案, 人工知能学会論文誌, 16 巻, 3 号 F (2001), 353–362.

付 録

1. 補題 1 の証明

式 (3) に従って得られる観測値 x_t に対し、式 (2) により求めた推定値 \hat{x}_t は以下のような式になる。

$$\begin{aligned}\hat{x}_{t+1} &= \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} x_{t-\tau} \\ &= \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} s_t - \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} \sum_{\tau'=1}^{\tau} v_{t-\tau'} \\ &\quad + \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} \epsilon_{t-\tau}\end{aligned}$$

この右辺第 2 項を τ' で整理すれば、

$$\hat{x}_{t+1} = s_t - \sum_{\tau'=1}^{\infty} (1-\alpha)^{\tau'} v_{t-\tau'} + \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} \epsilon_{t-\tau}$$

よって、予測誤差 δ_t は

$$\begin{aligned}\delta_t &= - \sum_{\tau'=1}^{\infty} (1-\alpha)^{\tau'} v_{t-1-\tau'} \\ &\quad + \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^{\tau} \epsilon_{t-1-\tau} - v_{t-2}\end{aligned}$$

となる。 ϵ_t , v_t は、各々の分散が σ_{ϵ}^2 , σ_v^2 で平均が 0 の独立な乱数なので、上記の誤差の二乗平均 $\mathbf{E}(\delta_t^2)$ は以下ようになる。

$$\mathbf{E}(\delta_t^2) = \frac{1}{2-\alpha} (2\sigma_{\epsilon}^2 + \frac{1}{\alpha} \sigma_v^2)$$

2. 定理 1 の証明

二乗誤差平均 $\mathbf{E}(\delta_t^2)$ の α についての偏微分を求めると、以下ようになる。

$$\begin{aligned}\frac{\partial \mathbf{E}(\delta_t^2)}{\partial \alpha} &= \frac{1}{(2-\alpha)^2} (2\sigma_{\epsilon}^2 + \frac{1}{\alpha} \sigma_v^2) + \frac{1}{2-\alpha} (-\frac{1}{\alpha^2} \sigma_v^2) \\ &= \frac{2(\alpha^2 \sigma_{\epsilon}^2 + (\alpha-1) \sigma_v^2)}{\alpha^2 (2-\alpha)^2}\end{aligned}$$

上式を 0 とおき、 $0 < \alpha < 1$ の範囲で解を求めれば、下記のようにになる。

$$\begin{aligned}\alpha &= \frac{-\sigma_v^2 + \sqrt{\sigma_v^4 + 4\sigma_{\epsilon}^2 \sigma_v^2}}{2\sigma_{\epsilon}^2} \\ &= \frac{-\gamma^2 + \sqrt{\gamma^4 + 4\gamma^2}}{2}\end{aligned}$$

3. 補題 2 の証明

まず、以下の補題を示す。
[補題 3]

$$\xi_{t+1}^{(k)} = \alpha^2 \sum_{\tau=0}^{\infty} \tau (1-\alpha)^{\tau-1} \xi_{t-\tau}^{(k-2)} \quad (\text{A-1})$$

□

証明

$$\begin{aligned} \eta_{t+1} &= \alpha^2 \sum_{\tau=0}^{\infty} \tau(1-\alpha)^{\tau-1} \xi_{t-\tau}^{(k-2)} \\ &= \alpha^2 \left[1(1-\alpha)^0 \xi_{t-1}^{(k-2)} + 2(1-\alpha)^1 \xi_{t-2}^{(k-2)} \right. \\ &\quad \left. + 3(1-\alpha)^2 \xi_{t-3}^{(k-2)} + \dots \right] \end{aligned}$$

と置くと、

$$(1-\alpha)\eta_t = \alpha^2 \left[1(1-\alpha)^1 \xi_{t-2}^{(k-2)} + 2(1-\alpha)^2 \xi_{t-3}^{(k-2)} \right. \\ \left. + 3(1-\alpha)^3 \xi_{t-4}^{(k-2)} + \dots \right]$$

よって、

$$\begin{aligned} \eta_{t+1} - (1-\alpha)\eta_t &= \alpha^2 \sum_{\tau=0}^{\infty} (1-\alpha)^\tau \xi_{t-1-\tau}^{(k-2)} \\ &= \alpha \xi_t^{(k-1)} \end{aligned}$$

これにより、以下の漸化式を得る。

$$\eta_{t+1} = (1-\alpha)\eta_t + \alpha \xi_t^{(k-1)}$$

これは $\xi_t^{(k)}$ の漸化式と同じである。よって、 $\eta_0 = \xi_0^{(k)}$ であれば、 $\eta_t = \xi_t^{(k)}$ となる。

これらにより、式 (A.1) が成立する。

この補題をもとに、補題 2 は以下のように導出できる。

まず $k=1$ の時考えると、

$$\begin{aligned} \frac{\partial \xi_t^{(1)}}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left[\alpha \sum_{\tau=0}^{\infty} (1-\alpha)^\tau x_{t-\tau-1} \right] \\ &= \sum_{\tau=0}^{\infty} (1-\alpha)^\tau x_{t-\tau-1} \\ &\quad + \alpha \sum_{\tau=0}^{\infty} (-1) \tau (1-\alpha)^{\tau-1} x_{t-\tau-1} \\ &= \frac{1}{\alpha} (\xi_t^{(1)} - \xi_t^{(2)}) \end{aligned}$$

よって、 $k=1$ において式 (7) が成立する。

また k 未満において式 (7) が成立しているとする、

$$\begin{aligned} \frac{\partial \xi_t^{(k)}}{\partial \alpha} &= \frac{\partial}{\partial \alpha} \left[\alpha \sum_{\tau=0}^{\infty} (1-\alpha)^\tau \xi_{t-\tau-1}^{(k-1)} \right] \\ &= \sum_{\tau=0}^{\infty} (1-\alpha)^\tau \xi_{t-\tau-1}^{(k-1)} \\ &\quad - \alpha \sum_{\tau=0}^{\infty} \tau (1-\alpha)^{\tau-1} \xi_{t-\tau-1}^{(k-1)} \\ &\quad + \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^\tau \frac{\partial}{\partial \alpha} \xi_{t-\tau-1}^{(k-1)} \\ &= \frac{1}{\alpha} \xi_t^{(k)} - \frac{1}{\alpha} \xi_t^{(k+1)} \\ &\quad + \alpha \sum_{\tau=0}^{\infty} (1-\alpha)^\tau \frac{k-1}{\alpha} (\xi_{t-\tau-1}^{(k-1)} - \xi_{t-\tau-1}^{(k)}) \\ &= \frac{1}{\alpha} \xi_t^{(k)} - \frac{1}{\alpha} \xi_t^{(k+1)} \\ &\quad + (k-1) \frac{1}{\alpha} \xi_t^{(k)} - (k-1) \frac{1}{\alpha} \xi_t^{(k+1)} \\ &= \frac{k}{\alpha} (\xi_t^{(k)} - \xi_t^{(k+1)}) \end{aligned}$$

よって、一般の k について式 (7) が成立する。 ■

4. 定理 2 の証明

$k=1$ の時、

$$\frac{\partial \bar{x}_t}{\partial \alpha} = \frac{\partial}{\partial \alpha} \xi_t^{(1)} = (-\alpha)^{-1} (\xi_t^{(1)} - \xi_t^{(2)})$$

よって、 $k=1$ で式 (8) が成立する。

また、 k 未満までで式 (8) が成立しているとする、

$$\begin{aligned} \frac{\partial^k \bar{x}_t}{\partial \alpha^k} &= -(k-1)(-1)^{-(k-1)} \alpha^{-k} (k-1)! (\xi_t^{(k)} - \xi_t^{(k-1)}) \\ &\quad + (-1)^{-(k-1)} \alpha^{-(k-1)} \left[\frac{\partial}{\partial \alpha} \xi_t^{(k)} - \frac{\partial}{\partial \alpha} \xi_t^{(k-1)} \right] \end{aligned}$$

この右辺の [] 内の第 1 項、第 2 項は各々 $\frac{k}{\alpha} (\xi_t^{(k)} - \xi_{t+1}^{(k)})$ 、 $\frac{k-1}{\alpha} (\xi_t^{(k-1)} - \xi_{t+1}^{(k-1)})$ なので、

$$\begin{aligned} \frac{\partial^k \bar{x}_t}{\partial \alpha^k} &= (-1)^{-(k-1)} \alpha^{-k} (k-1)! \\ &\quad \times \left[-k \xi_t^{(k+1)} + (k + (k-1) - (k-1)) \xi_t^{(k)} \right. \\ &\quad \left. + ((k-1) - (k-1)) \xi_t^{(k-1)} \right] \\ &= (-1)^{-(k-1)} \alpha^{-k} (k-1)! \left[-k \xi_t^{(k+1)} + k \xi_t^{(k)} \right] \\ &= (-\alpha)^{-k} k! (\xi_t^{(k+1)} - \xi_t^{(k)}) \end{aligned}$$

よって、一般の k において式 (8) が成立する。 ■

なお、一般の REMA $\xi_t^{(k)}$ についても、同様の帰納法を用いて、 α による m 階偏微分を以下のように導出できる。

$$\frac{\partial^m \xi_t^{(k)}}{\partial \alpha^m} = \frac{k}{\alpha^m} \sum_{i=0}^m (-1)^i \frac{m!}{i!(m-i)!} \frac{(k+i-1)!}{(k+i-m)!} \xi_t^{(k+i)}$$