

# AR モデル

岡田 大 (Okada Masaru)

October 8, 2025

## Abstract

線形な時系列データの分析に用いられる AR モデルを概観する。

## 1 線形予測モデル

線形予測モデルには様々あり、

1. AR (Auto Regression) モデル
2. MA (Moving Average) モデル
3. AR と MA を組み合わせた ARMA (Autoregressive Moving Average) モデル
4. 自己階差を入れた ARIMA (Autoregressive Integrated Moving Average) モデル

等がある。

これらの線形予測モデルは線形の範囲内で時系列データの解析には有用である一方で、デメリットとしてはカオス等を含む非線形ダイナミクスを扱うことはできない。

今回は最も基本的な線形予測モデルの一つである AR モデルについてノートにまとめた。

## 2 確率過程と決定論的過程

$N$  個の観測値から構成される時系列を

$$\{x(t)\}_0^{N-1} = \{x(t_0), x(t_0 + \Delta t), x(t_0 + 2\Delta t), \dots, x(t_0 + (N-1)\Delta t)\}$$

のように表す。 $t$  は 0 から  $N-1$  までの整数値を取る。

$\Delta t$  は隣り合う観測値間の時間間隔を表す。

$t_0$  は観測を始める初期時刻であり、 $t_0 + (N-1)\Delta t$  に観測が終了したデータになっている。

つまり、実現可能な挙動のごく一部しか見ていない。 $\{x(t)\}_0^{N-1}$  は一つのサンプルに過ぎない。

そのため、 $\{x(t)\}_0^{N-1}$  を分析して時系列推定をしても、真の性質から離れたものである可能性は常にある。

推定結果が真の性質にどれほど近いかを評価するのも一つの大きな問題である。

$\{x(t)\}_0^{N-1}$  を構成する各データを、ある確率で実現された変数  $X$  の値、すなわち、確率変数の実現値と考える。

確率によってモデル化された時間発展のダイナミクスのことを確率過程という。

特に、各時刻での実現確率が 1 に等しい場合は決定論的であるという。

実在するシステムの状態変化は、物理法則や何らかの因果則に従って起こる。

その意味で、時間変化は本来決定論的であるはずで、本質的に因果的でない過程は量子力学における波束の収縮過程だけであろう。

しかし状態変化が非常に多くの要因によって決まっていたり、外部からの制御不能な影響のために内部要因を正確に設定することが実際的に不可能な場合がままある。

こういったケースでは時間発展のダイナミクスをデータに基づいて分析する上で、知識に不足が生じる。知識の不足に応じて、観測される挙動は決定論的過程から遠ざかっていくように見える。

### 3 定常状態

定常確率過程、すなわち確率過程が定常状態であるとは、時系列が上昇し続けたり、下降し続けたりせず、一定の水準付近に留まっているような、統計的な平衡状態にあることを意味する。

確率密度関数が時間に依存しない場合、系は定常確率過程となる。

すなわち、ある時刻  $t_0$  から始まる時系列  $\{x(t_0)\}_0^{N-1}$  と、時刻  $T$  だけ遅れて始まる時系列  $\{x(t_0 - T)\}_0^{N-1}$  とが、同一の結合分布確率を持つときに（強）定常過程と呼ぶ。

$$F[\{x(t_0)\}_0^{N-1}] = F[\{x(t_0 - T)\}_0^{N-1}]$$

### 4 線形ダイナミクス

以下では記号を簡単にするための約束として、 $\{x(t)\}_0^{N-1}$  の各要素は平均  $\mu$  はゼロとする。つまり処理としてすでに以下が施されているものとする。

$$x(t) \leftarrow x(t) - \mu$$

観測値の列からベクトルを構成する。

$$\mathbf{x}(t) = (x(t), x(t-1), x(t-2), \dots, x(t-(D-1)))$$

$D = \infty$  は無限の過去までさかのぼることを意味する。

現在を起点にして  $\tau$  ステップだけ未来における挙動  $x(t + \tau)$  が、状態変化を表す写像  $F$  を用いて、

$$x(t + \tau) = F[\mathbf{x}(t)]$$

のように表されるものとする。

この  $F$  について、

$$F(a\mathbf{x} + b\mathbf{y}) = aF(\mathbf{x}) + bF(\mathbf{y})$$

が満たされるのであれば  $F$  は線形ダイナミクスを表す。これが我々の分析対象とするダイナミクスである。

$F$  が線形ダイナミクスでないとき、非線形ダイナミクスという。これについてもまた別のノートにまとめたい。

### 5 Wold の分解定理

どのような定常過程  $\{z(t)\}$  も、決定論的な定常過程  $\{y(t)\}$  と非決定論的な定常過程  $\{x(t)\}$  の和で表すことができる [1]。

$$z(t) = x(t) + y(t)$$

これを Wold の分解定理と呼ぶ。

興味があるのは非決定論的な定常過程  $x(t)$  である。

非決定論的な定常過程は、平均値がゼロで、分散が 1 のホワイトノイズ  $\xi(t)$  の無限級数で与えられる。

$$x(t) = \xi(t) + \sum_{i=0}^{\infty} a_i \xi(t-i)$$

ただし、この無限級数の係数は絶対収束する。

$$\sum_{i=0}^{\infty} |a_i| < \infty$$

双対性を用いて以下のように書き直すこともできる。

$$x(t) = \xi(t) + \sum_{i=0}^{\infty} c_i x(t-i)$$

この無限級数の係数も絶対収束する。

$$\sum_{i=0}^{\infty} |c_i| < \infty$$

いずれの方程式も無限級数を含むので、これらを直接用いて時系列予測することはできない。

そこで、有限項で打ち切って確率過程を近似することを考える。これが AR 過程になる。

## 6 自己回帰 (AR) モデル

自己回帰過程 (AR 過程) とは

$$x(t) = \xi(t) + c_1 x(t-1) + c_2 x(t-2) + \cdots + c_p x(t-p)$$

によって表される過程である。

上の式のように  $p$  個の係数を含む過程は  $p$  次 AR 過程と呼ばれる。この  $p$  個の係数は AR パラメータと呼ばれる。 $p$  個の AR パラメータによって近似されるモデルを AR( $p$ ) モデルと呼ぶ。

特に、AR(1) 過程で係数が 1 に近いものを単位根 (unit root) と呼ぶ。

## 7 AR モデルの推定

### 7.1 Yule-Walker 方程式

AR( $p$ ) 過程の式

$$x(t) = \xi(t) + c_1 x(t-1) + c_2 x(t-2) + \cdots + c_p x(t-p)$$

この両辺に  $x(t-\tau)$  を乗じると、

$$x(t)x(t-\tau) = \xi(t)x(t-\tau) + c_1 x(t-1)x(t-\tau) + \cdots + c_p x(t-p)x(t-\tau)$$

自己共分散

$$\gamma(\tau) = E[x(t)x(t-\tau)]$$

を定義して、 $\tau = 0$  で両辺の期待値  $E[\cdot]$  を取ると、

$$x(t)^2 = \sigma + c_1 x(t-1)x(t) + c_2 x(t-2)x(t) + \cdots + c_p x(t-p)x(t)$$

より、

$$\gamma(0) = c_1 \gamma(1) + c_2 \gamma(2) + \cdots + c_p \gamma(p)$$

ホワイトノイズの分散は  $E[\xi(t)x(t)] = \sigma$  と置いている。

また、 $E[\xi(t)x(t-\tau)] = 0$  ( $\tau > 0$ ) であることを用いると、

$$\gamma(\tau) = c_1 \gamma(\tau-1) + c_2 \gamma(\tau-2) + \cdots + c_p \gamma(\tau-p)$$

この  $\tau$  に  $1, 2, \dots, p$  を代入することで、Yule-Walker 方程式と呼ばれる連立方程式が得られる。

$$\sigma = \gamma(0) - c_1 \gamma(1) - c_2 \gamma(2) - \cdots - c_p \gamma(p)$$

$$\gamma(1) = c_1 \gamma(0) + c_2 \gamma(1) + c_3 \gamma(2) + \cdots + c_p \gamma(p-1)$$

$$\gamma(2) = c_1 \gamma(1) + c_2 \gamma(0) + c_3 \gamma(1) + \cdots + c_p \gamma(p-2)$$

$$\vdots$$

$$\gamma(p) = c_1 \gamma(p-1) + c_2 \gamma(p-2) + c_3 \gamma(p-3) + \cdots + c_p \gamma(0)$$

ただし、 $\gamma(k) = \gamma(-k)$  の対称性を用いている。

この連立方程式を解いて AR パラメータを求めることで AR モデルを決定できる。

AR パラメータを求めることで線形予測子

$$\hat{x}(t) = c_1 x(t-1) + c_2 x(t-2) + \cdots + c_p x(t-p)$$

の最良近似を実現する。

線形予測子の予測誤差を

$$H_{\text{lin}} = \frac{1}{N} \sum_{t=p}^{n-1} \left( x(t) - \hat{x}(t) \right)^2$$

と置く。 $H_{\text{lin}}$  を最小にする AR 係数は

$$\frac{\partial H_{\text{lin}}}{\partial c_i} = 0$$

から得られるが、これは Yule-Walker 方程式を解いて得られる係数と一致する。

次数  $p$  が高くなる場合は Yule-Walker 方程式を高速で解くアルゴリズムが必要になる。

Levinson-Durbin のアルゴリズムなどが高速なアルゴリズムとして知られている [4], [5]。

## References

- [1] Box, G.E.P., Jenkins, G.M. and Reinsel, G.C. (1994) Time Series Analysis; Forecasting and Control. 3rd Edition, Prentice Hall, Englewood Cliff, New Jersey.

- [2] Yule, G. Udny (1927), “On a Method of Investigating Periodicities in Disturbed Series, with Special Reference to Wolfer’s Sunspot Numbers” , Philosophical Transactions of the Royal Society of London, Ser. A 226: 267 – 298
- [3] Walker, Gilbert (1931), “On Periodicity in Series of Related Terms” , Proceedings of the Royal Society of London, Ser. A 131: 518 – 532 Hamilton (1994), p. 59
- [4] Levinson, N. (1947). ”The Wiener RMS error criterion in filter design and prediction.” J. Math. Phys., v. 25, pp. 261 – 278.
- [5] Durbin, J. (1960). ”The fitting of time series models.” Rev. Inst. Int. Stat., v. 28, pp. 233 – 243.