

ベイジアンネットワーク

1. 概要

ベイジアンネットワークとは過去の事象から現在の事象に影響を与えるというベイズ推定に基づく確率的なグラフィティカルモデルである。モデル選択基準によるベイジアンネットワークの構造推定を行ったのち、そのベイジアンネットワークのモデルで予測するという手順を踏む。ベイジアンネットワークの構造の探索モデルとして、全探索法、MWSTアルゴリズム、K2アルゴリズム、遺伝アルゴリズムによる探索だ。しかし、データに時系列性があればさらに探索モデルを短縮できるのではないかと考えた。

1.1 時系列分析手法

時間経過ごとに記録された数値列からモデルを作成して、将来の予測を行う分析手法のことである。

1.1.1 ARモデル

AR(p)モデルでは、t期の株価リターンを r_t をp期前までおリターンと、攪乱項 u_t によって表す。

$$r_t = \alpha_0 + \sum_{i=1}^p \alpha_i r_{t-i} + u_t$$

ここで、 α_i はモデルのパラメータである。 u_t は平均 0、分散 σ^2 のホワイトノイズに従う。

2. ベイジアンネットワーク

2.1 ベイジアンネットワークの探索方法

探索アルゴリズムに関して、構造候補数の大きさについて考える。構造候補数の大きさから、それぞれの探索アルゴリズムがどういう関係にあるかを考える。

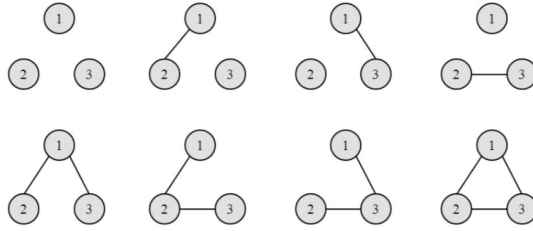
2.1.1 遺伝アルゴリズム

植野 (Ueno,1994)、ラー(Larran~aga、1996a,b)、ウォン(Wong,1999)らは、全順序関係を制約としない探索として、遺伝アルゴリズムを提案している。遺伝アルゴリズムの構造は、ノードの数をnと置くと、 $S = \{s_{ij}\}$ とする正方行列で考えらえる。 s_{ij} は変数iから変数jにアークがあれば1,そうでなければ0の値をとる。よって、これの構造候補数の式は

$$F(n) = 2^{2n}$$

真の構造を見つけるのが困難な大規模なネットワーク構造の同定では、非常に有効であることが報告されている。

2.1.2 全探索アルゴリズム



上の図では、3つの確率変数ノードで無向なグラフを考えている。ノード3つを考えると、矢の方向がどちらにも向くことを考えれば

$$F(n) = \sum_{i=1}^n (-1)^{i+1} (nCi) 2^{i(n-i)} F(n-i)$$

で与えられることが知られている。

2.1.3 K2アルゴリズム (欲張り法)



上の図のように全順序関係が知られているとする。ノードの全順序関係を知っている有向非循環グラフと無向グラフの構造候補数は同じだと考えられるので、式は以下ようになる。

$$F(n) = F(n-1)2^{n-1} = 2^{\sum_{i=1}^{n-1} i}$$

となる。しかし、ここでは全順序関係はあるが、時系列データであると限定していない。

2.1.4 MWST法

チョウ(Chow, 1968)によると相互情報量による木の生成手法が提案されている。n個の離散変数集合 $u = \{X_1, X_2, \dots, X_n\}$ について考える。真の同時確率分布、 $P_0 = p(X_1, \dots, X_n)$ と $P_1 = p(X_1, \dots, X_n | B_S)$ の構造を考え、2つの分布の距離をカルバックライブラー測度によって

$$D(P_0, P_1) = \sum_{X_1, \dots, X_n} P_0 \log \frac{P_0}{P_1} \geq 0$$

と定義できる。ここでDを最小化するような P_1 を見つければ良い。このとき、以下の定理が知られている。

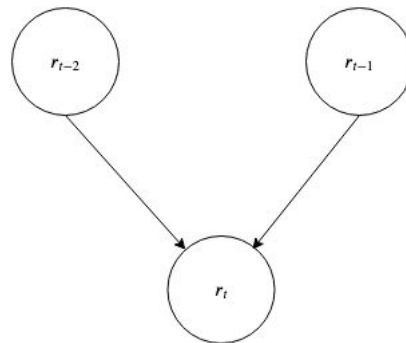
$D(P_0, P_1)$ を最小化する P_1 は、構造 B_S における全ての枝の相互情報量 $I(X_i, X_{i'})$ の和を最大にする。ここで X_i と $X_{i'}$ はネットワーク中の任意のノード変数を示している。構造候補数に関しては

$$F(n) = \frac{n(n-1)}{2}$$

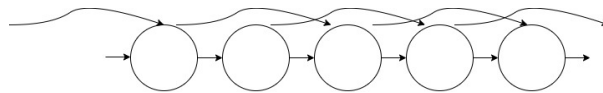
として与えられることが考えられる。

2.1.5 提案アルゴリズム

時系列性データであると限定した場合、K2アルゴリズムを短縮して計算量を減らして計算できるのではないかと考えた。

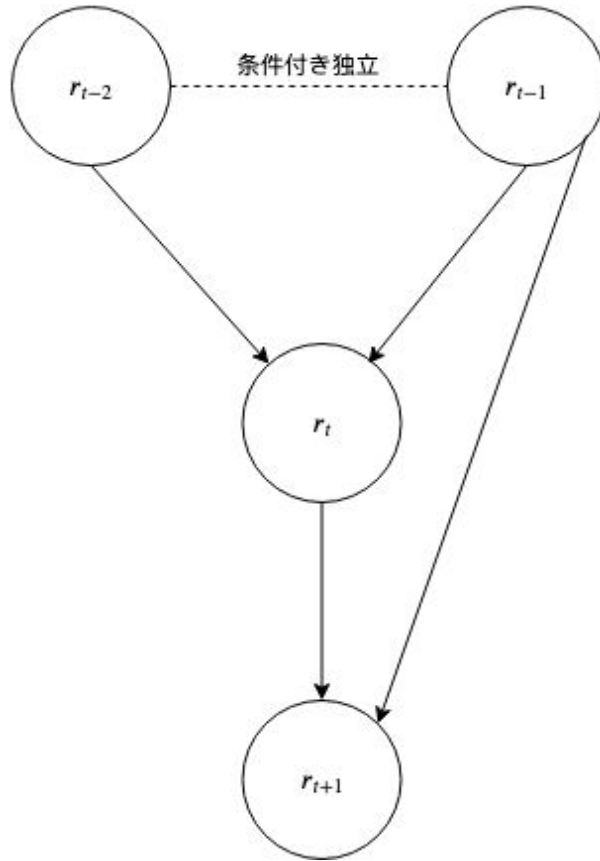


ノードが3つとして仮定する。 $\{r_{t-1}, r_{t-2}\}$ を観測値とし、 $\{r_t\}$ を予測値と考えてベイジアンネットワークを決定すれば良い。仮に上の図のようにベイジアンネットワークが決定されるならば、時系列データは以下のような図で表せる。



よって、一度 r_t を予測するためのベイジアンネットワークを決定すれば、 r_{t+1} 以降のデータも予測することが可能である。

上の例で考えると、 $P(r_t, r_{t-1}, r_{t-2}) = P(r_t | r_{t-1}, r_{t-2})$ として同時確率分布が与えられる。



また、 r_{t+1} についても、 $P(r_{t+1}, r_t, r_{t-1}) = P(r_{t+1}|r_t, r_{t-1}) = P(r_{t+1}|r_{t-1})P(r_t|r_{t-1}, r_{t-2})$ のように同じく求まる。この手法で考えると、構造候補数を

$$F(n) = 2^{(n-1)}$$

と短縮して考えられる。

2.2 ベイジアンネットワークでの予測

ベイジアンネットワークで予測するステップは以下ようになる。

Step1. 株価リターンを算出する。

Step2. 株価リターンを離散化する。

Step3. 離散化されたデータを用いて、2.1.5で提案した手法によりBNを決定する。

Step4. BNによって予測された確率により期待値を算出する。

2.2.1 株価リターンの算出

ある株価データの終値を対数差分で表す。

$$r_t = \log \frac{P_t}{P_{t-1}} \times 100$$

P_t は t 期の株価終値を表す。

2.2.2 株価リターンの分割

株価収益率をL個にクラスタリングする。各クラス C^l の重心を r^l と置く。離散値集合は以下のようなになる。

$$\{r^1, \dots, r^L\}$$

また本稿では、株価リターンにおける等分割クラスタリングとワード法によるクラスタリングと、株価リターンと出来高の2つの変数におけるワード法でのクラスタリングを用いる。

2.2.3 ベイジアンネットワークの決定

2.1.5で提案した手法により計算し、AICによりモデルを選択する。AICは以下のように計算される。

$$AIC = -2l_m(\theta_m|X) + 2k_m$$
$$l_m(\theta_m|X) \sim \sum_{j=1}^q \sum_{k=1}^r (N_{jr}) \log \theta_{jr}$$
$$k_m = q$$

により情報量基準を求める。

2.2.4 株価リターンの予測

予測においては期待値による算出を行う。

$$r_t = r^l \times P(r^l|B_S)$$

3 予測の精度について

予測の精度をはかる指標としてRMSEを用いる。

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (r_t - r_t')^2}$$

実際の株価リターンを r_t 、予測の株価リターンを r_t' と定義する。

参考文献

繁榊算男・本村陽一, 植野真臣 : ベイジアンネットワーク概説, 培風館, 2006

植野真臣 : ベイジアンネットワーク, コロナ社, 2013

左毅 : ベイジアンネットワークの株価予測への応用, 情報処理学会論文誌数理モデル化と応用 (TOM) , 4(4), 92-103, 2011-11-30

P.Larranaga, C.M.H.Kuijpers, R.H.Murga, Y.Yurramendi : Learning Bayesian Network Structures by Searching For the Best Ordering With Genetic Algorithms, IEEE Transactions on Systems, Man and Cybernetics, Vol.26, No.4, pp.487-493, 1996

J.A.Hartigan and M.A.Wong : A K-Means Clustering Algorithm, Applied Statistics, Vol.28, No.1, pp.100-108, 1979

M.L.Wong, W.Lam,K.S.Leung : Using Evolutionary Programming and Minimum Description Length Principle for Data Mining of Bayesian Networks, IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.21, No.2, pp.174-178, 1999

Chow, C.K., and Liu, C.N. : Approximating discrete probability distributions with dependence trees. IEEE Transactions on Information Theory, IT-14, pp462-467, 1968