楕円の相関係数（その１０）

nino

2017 年 9 月 10 日

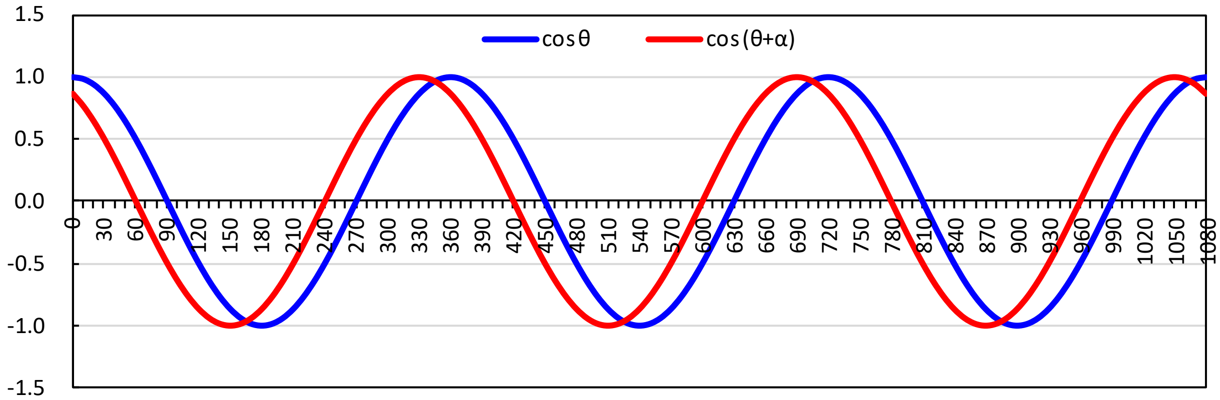
最後に，相関係数の応用例のひとつとして，気温と太陽高度や水温の季節変化など同一周期の時系列データの位相差を相関係数を用いて求める方法について考察します．

周期時系列のコサイン関数モデル

相関係数は，２つの三角関数の位相差のコサインで表されました．言い換えると，２変数の相関係数のアークコサインは２変数の位相差を示しています．２変数が周期時系列データであれば，その相関係数から位相差を求めることができるわけです．相関係数をR，位相差をαとすると，α＝arccos(R)が成り立ちます．

図には，２つの周期時系列を$x_1 = \cos\theta$，$x_2 = \cos(\theta + \alpha)$とし，α＝30度の場合における２変数の３周期の時系列を示しました．横軸の単位は「度」です．

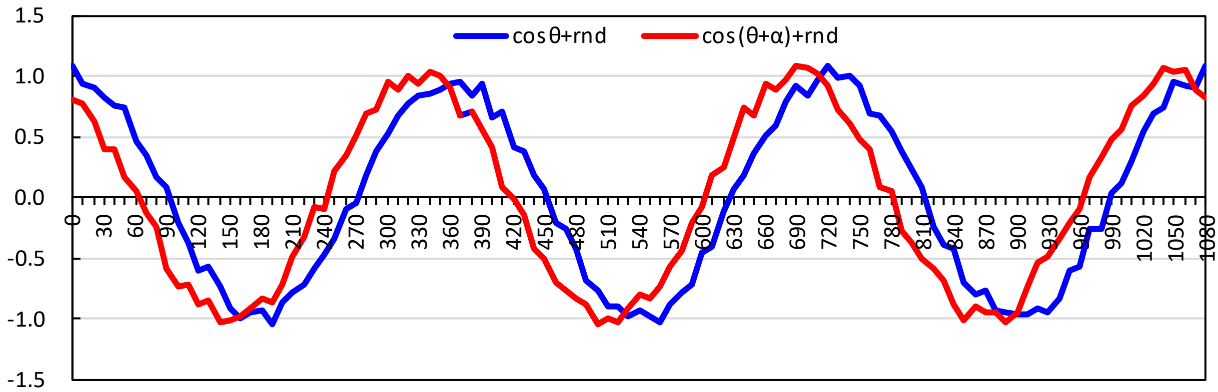
２つの変数x_1とx_2の相関係数は0.866で，そのアークコサイン（位相差）は理論値の30.00度となります．なお，位相差は相互相関係数などを用いて求める方法もあります．



ノイズを含むコサイン関数モデル

実際の測定値は，測定誤差や外乱などの影響を受けており，コサイン関数のような時系列変化を示すことは少ないと思います．次図は，先の２変数にノイズとして－0.1～0.1の範囲（±10%）で発生させた乱数(rnd)を加えたもので，２変数を$x_1' = \cos\theta + \mathrm{rnd}$，$x_2' = \cos(\theta + \alpha) + \mathrm{rnd}$で表しています．

このノイズを付加させた２つの変数の相関係数は0.856で，そのアークコサイン（位相差）は31.11度となりました．このように，ノイズが加わると相関係数は低下し，その結果として位相は理論値より大きくなります．ただし，ノイズを含む場合でも概略の位相差を求められますので，複数変数の実測値の相関係数表を位相差表に変換することによって，それら相互の位相差を知ることができます．なお，実測値による位相差の偏りを小さくし，より正確な値を求めるには，最小二乗法を用いる方法があります．



最小二乗法による実測値のコサイン関数モデルへの近似

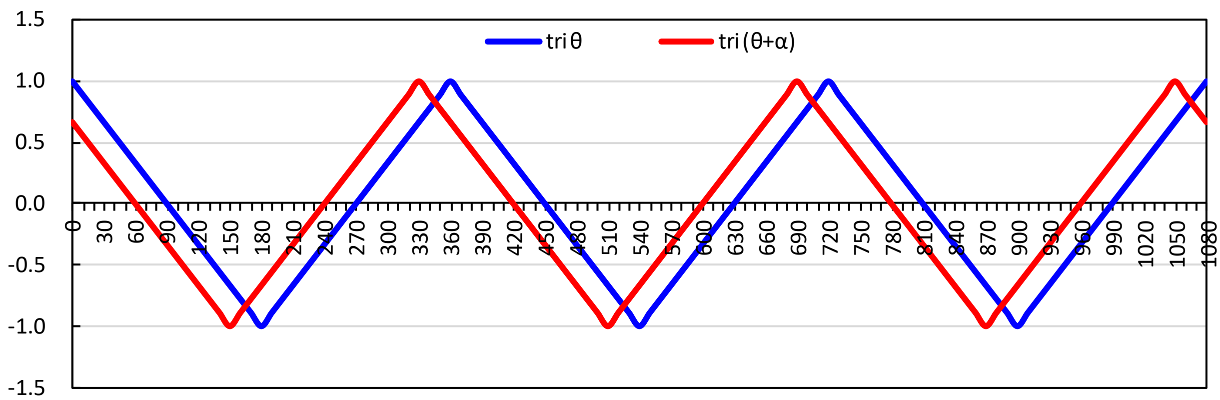
偏りを小さくする手段のひとつとして，最小二乗法を用いて実測値をコサイン関数に近似する方法が考えられます．具体的には，まず，２つの変数それぞれについて，測定時間ごとの実測値とコサイン関数の差の二乗を計算し，次に，その二乗和が最小になるコサイン関数モデルのパラメーター（振幅と位相）をExcelのソルバーアドインにより求めて，最後に，２つのコサイン関数モデルの相関係数のアークコサインから位相差を求めます．なお，この値は，パラメーターとして得られた２つのモデルの位相の差と一致します．

ソルバーアドインについては，（その７）を参照してください．この例では，最小となる数式は，実測値とコサイン関数モデルの差の二乗和のセルであり，変化する数値は，モデルのパラメーター（振幅と位相）です．

ノイズを含むコサイン関数モデルのパラメーターを計算すると，変数x_1'の振幅と位相は1.01，-0.40度，変数x_2'の振幅と位相は0.98，29.99度でした．２変数x_1'とx_2'の相関係数は0.863，そのアークコサインは30.39度となり，ノイズを含むコサイン関数より理論値に近い値が得られました．なお，得られたパラメーターの位相から，位相差は29.99－(－0.40)＝30.39度となり，相関係数から求めた値と一致します．

三角波の相関係数と位相差

実際の測定値の時系列は，コサイン関数のような変化を示さない場合が多いと思います．その簡単な例として，波形が三角波（略号：tri）の時系列について同様に検討します．

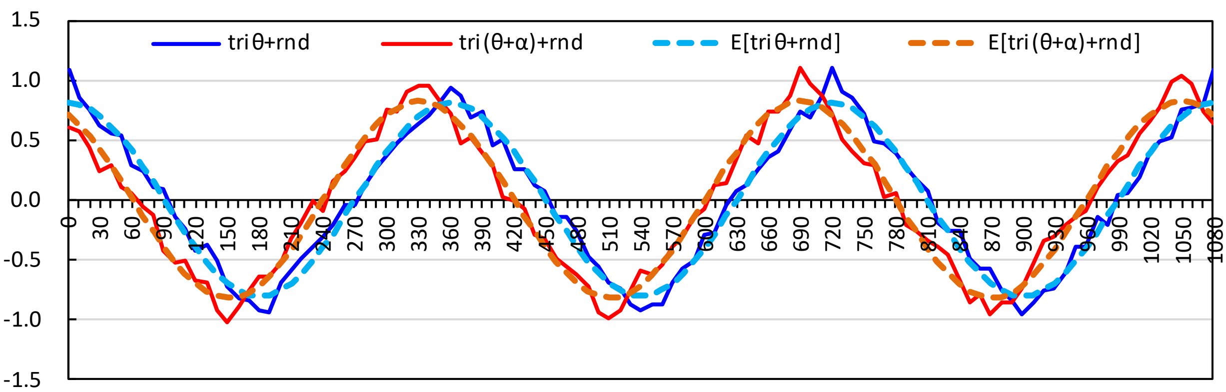


三角波の振幅と位相差は，前述のコサイン関数と同じ1.00と30.00度です．それらの相関係数とアークコサインは，それぞれ0.851，31.71度となりました．三角波のため，コサイン関数に比べて，相関係数は小さく，位相差は大きくなっています．

三角波を前述した最小二乗法を用いてコサイン関数モデルに近似し，そのモデルの相関係数から位相差を求めてみました．最小二乗法を適用した２つのモデルの相関係数は0.866，位相差は30.00となり，理論値を一致しました．なお，推定されたパラメーターの振幅は0.80で，1.00より小さい値を示します．

ノイズを含む三角波の最小二乗法によるコサイン関数モデルへの近似

ノイズを含む三角波について，同様に計算してみます．ノイズは，前述のコサイン関数の場合と同様に，－0.1～0.1の範囲で発生させた乱数(rnd)です．
その相関係数とアークコサインは，それぞれ0.848，32.04度となり，ノイズが加わった分だけノイズ無しの三角波よりも，相関係数は小さく，位相差は大きくなっています．



ノイズを含む三角波を最小二乗法によりコサイン関数モデルに近似した結果，2つのモデル(E[triθ+rnd]とE[tri(θ+α)+rnd])のパラメーター：振幅と位相はそれぞれ0.81と－0.14，0.82と29.40度でした．そして，２変数の相関係数は0.870，そのアークコサインは29.54度となり，それぞれノイズを含む三角波の値に比べて改善され，より理論値に近い値が得られました．

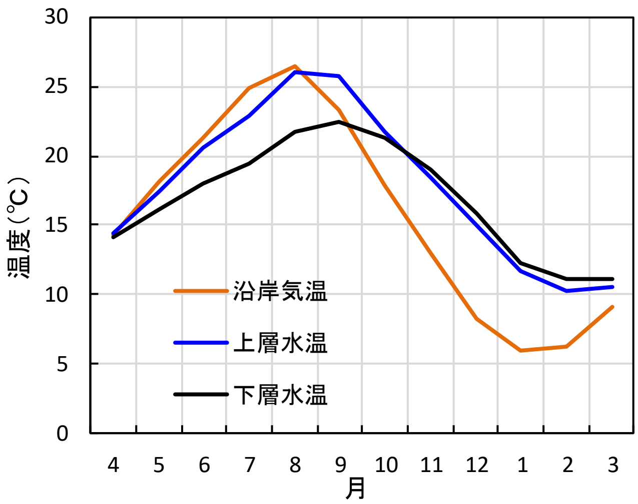
気温と水温の位相差

この方法を適用した簡単な具体例を紹介します．
図は，２０年間（1985～2004年度）における東京湾沿岸８地点の気温と東京湾内９地点の上層および下層の水温の実測値の月別平均値です．
最高温度は沿岸気温に対して上層水温は少し遅れ，下層水温はさらに遅れています．海水の熱容量や熱伝導率そして海流などが影響して，上下層水温は気温よりも季節変化が遅れて現れます．その遅れ時間を相関係数から求めてみます．

左表は３項目の相関係数です．右表は沿岸気温に対する上下層の位相差（単位：月）であ

り，それぞれの相関係数のアークコサイン(度)に12(月)／360(度)をかけた値です．
相関係数は，沿岸気温と上層水温および下層水温でそれぞれ0.931，0.837となりました．
それらの位相差はそれぞれ0.72月，1.11月となり，気温に対して上層水温は0.72月遅れ，
下層水温は1.11月遅れで最高温度は現れることがわかりました．

最小二乗法を用いて実測値をコサイン関数モデルに近似した場合は，相関係数は，沿岸気温と上層水温および下層水温でそれぞれ0.934，0.850，位相差（月）は，0.70月，1.06月となりました．位相差は0.1月程度の誤差で両者は一致しています．



[相関係数]	沿岸気温	上層水温	下層水温
沿岸気温	1.000	0.931	0.837
上層水温		1.000	0.973
下層水温			1.000

[位相差(月)]	沿岸気温	上層水温	下層水温
沿岸気温	0.00	0.71	1.11
上層水温		0.00	0.45
下層水温			0.00

このように，周期時系列をコサイン関数モデルに近似し，その相関係数から位相差を求める方法は比較的簡単です．しかし，実際の波形は様々なので，この手法をすべての周期時系列に適用できるわけではありません．波形とそのコサイン関数モデルの視認と周期や位相差の根拠などを充分に把握しておく必要があります．

また，周期時系列には同一周期の時系列だけではなく，異なる周期や複数の周期をもつ時系列もあります．これらの場合は，周期別にあるいは類似の周期特性をもつグループに群分けしてから，群別に位相差を求める必要があります．群分けには主成分分析やフーリエ解析などが有用です．