

確率と最小二乗法による土地の境界復元

準拠点選択

「交点計算の注意」「²検定・t検定」と合わせてご覧ください。

境界復元計算するには基準とする点が必要です、その基準とする点を準拠点といいます。理想的な方法は復元しようとする点に対して最も復元精度の高い点を準拠点とすることです。

境界復元の理想は正確に復元することと誰が復元しても大差ない位置に復元できることです。

さらに計算処理に時間が掛からないことが条件でしょうか。

測量するときに基準とする点を決める基準点測量が行われます、この時の基準点が現地に残っていればその基準点を使って境界は復元できます。

実際には多くの基準点は現地に残っていませんが境界点は残っていますので残っている境界点の中から基準にする点を選んで計算することを考えます。

点の中から基準にできる点をどのような考えでどのような方法で選ぶかを考えてみます。境界の復元と同じで準拠点選択でも正確性、普遍性、迅速性が求められます、迅速性に関しては大量にデータ計算を伴いますので今まで統計的な手法は避けたい原因はここにあります。

ウインドウズ95以降パソコンの普及に伴い2005年あたりからExcelを使った準拠点選択が簡単にできるようになりました。

当webでも2005年からExcel出作ったプログラムを公開していますので2005年より前に勉強されたかたと2005年以降に勉強された方で大きな差があるように感じます。

準拠点選択の意義とイメージ

正確性

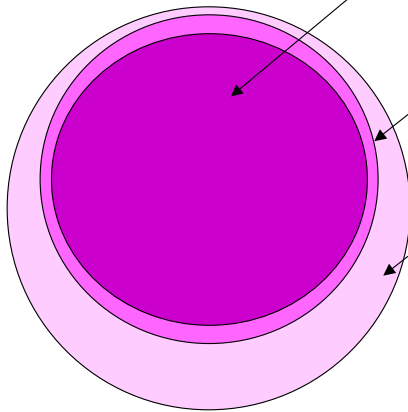
精度の高い復元値が得られること

普遍性

誰が復元しても変わらない値

迅速性

計算に時間が掛からないこと



不動点 (正規分布になる点)

正規分布になる点の組み合わせを
探す理由は「復元の基礎」「準拠点
選択の意義」で説明してあります

不動点 (正規分布から外れる点)

通常の測量であれば正規分布から
外れる点はない

初めから異常な点、異常な経
年変化のある点、

現地調査、聞き取りの時点で
除くこと

通常の測量であれば異常な点は無
い、境界標の管理がされていれば
異常な経年変化点はない

計算の基準になる点、これを準拠点と言いますがこれを選ぶ事が重要になってきます。

書証、物証、人証によって選ぶ、統計的検定で選ぶ、幾何学的観点から選ぶことが必要になります。

この選ぶことを準拠点選択といいまして端的に言えばその図面が描かれた時の点を特定する、その中からその時から動いていない点いわゆる不動点として考えられる点を探します。

この不動点と考えられる点からさらに確率論によって異常な点を除く、幾何学的な配点バランスを考慮して計算上使える不動点を選びます。

重要なのは図面が書かれた時の点から不動点を特定することで、書証、物証、人証の観点から現地を観察し関係者から状況を聞いて図面作成当時から動いていない点を計算の前に特定しておかなければなりません。

改祖図、更正図（明治に作られた公図）では境界に関する歴史的なことや地域の慣習等を考慮しなければなりません。

その後確率論によって異常な点を除くのは統計手法によって行いますのでさほど難しくはありません。

ここでは統計的手法についての解説を主に幾何学的な見方をベクトル図によって説明します。

準拠点数による復元精度

標準偏差、復元精度を変数減少法により計算する。

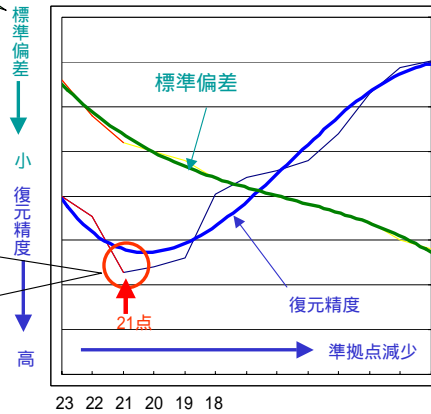
変数減少法

全変量から始めて、順次変数を減少させていく方法。標準偏差に悪い影響の大きい変数を順次取り除いていく方法。

標準偏差に関係なく復元精度の良い点の準拠点の組み合わせがある。

準拠点の選択とはこのポイントを探すことにつきます。

正規分布に近いデータ



の精度のよい点の範囲が狭いので簡単に見つかる

このデータは正解のある作られたデータから計算したもので実際のデータからは計算できません。

準拠点数における復元精度を変数減少法によって調べてみます。

変数減少法とは全変量から始めて、順次変数を減少させていく方法で標準偏差に悪い影響をを与えている大きい変数を順次取り除いていく方法です。

変数減少法によって点を除いていくと標準偏差はグラフのように小さくなります、このとき既知点の復元精度を計算しグラフにすると一旦、精度が高くなりその後精度が落ちていきます。

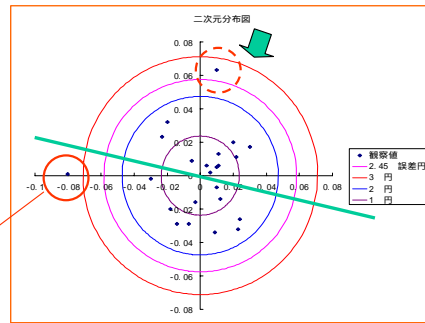
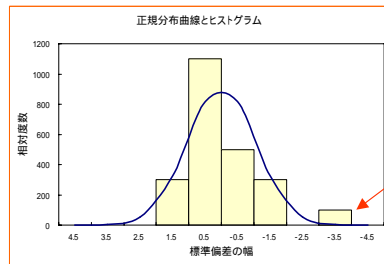
そこで精度の良くなる赤丸の点の組み合わせを探せばよいわけです。

次のスライドで説明しますがこの21点のところでは最初に正規分布になる点の組み合わせです。

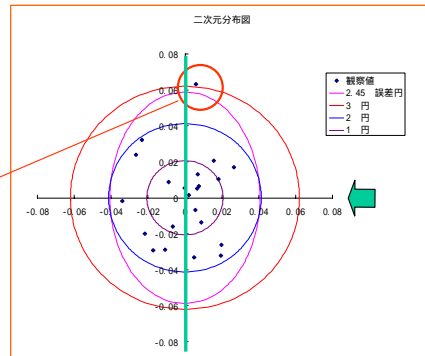
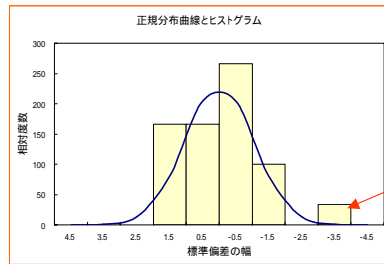
データにはそれぞれ個性があり一概にこれといった方法はなかなかありませんが概ね正規分布に最初になる点の組み合わせの位置が復元精度が高いことがわかります。

23点、22点 と異常点のある状態

23点



22点

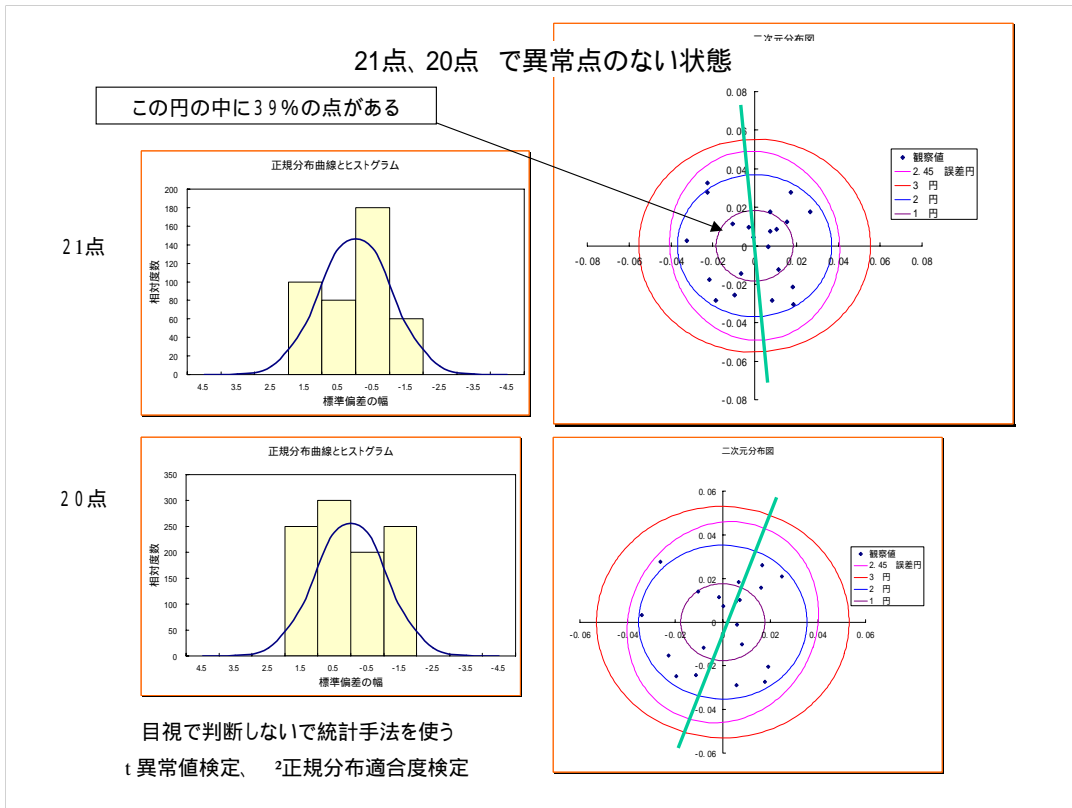


前のグラフの23点のデータです。

23点のデータを一点ずつ変数減少法で除いていったものです。

この場合では点数が少なくても大きく外れている点は容易に判断出来ます。

このような異常な点が明確に現れるデータが散布図からでも判断出来ます。



この図の状態になると果たしてこれが正規分布に倣うデータなのか否かは判断出来ないでしょう。

そこで統計手法の t 検定と χ^2 検定を使って判断することになります。

その結果、21点のところから χ^2 正規分布適合度検定で最初に正規分布になったと判断された点の組み合わせです。

検定の説明はこれだけで数時間かかりますので省略します、要するに統計的に判断するということを覚えておいてください。

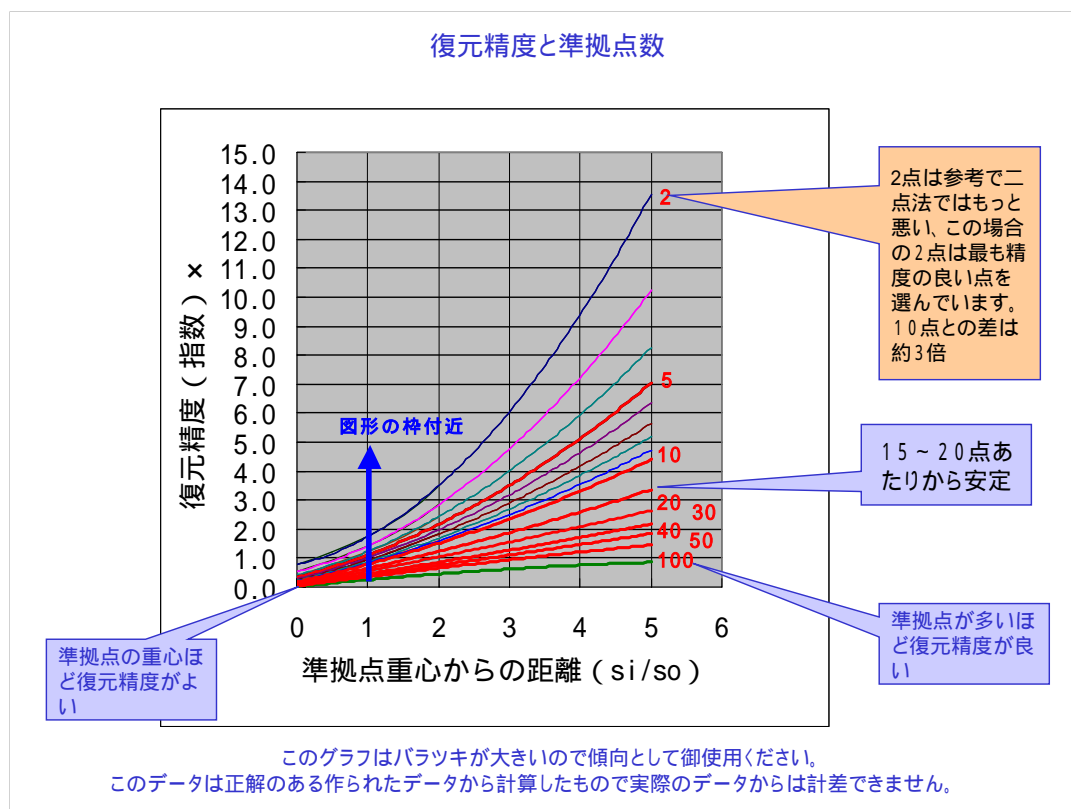
正規分布確率を使っても t 検定と χ^2 検定と近いデータが得られますが.....

2005年以前はこのような方法があっても計算が複雑で実用化されませんでした。パソコンの普及に伴い誰でもできるようになりました。

この例では21点、20点とも正規分布に従っています、この場合点数の多い21点の組み合わせを採用します。

その理由は次の通りです。

復元精度と準拠点数



このグラフは復元点に対して準拠点数、準拠点重心からの距離と復元精度の関係を調べた図です。

縦に復元精度、下に来るほど復元精度が良くなっています。

横に準拠点重心からの距離をとり、この関係を準拠点数毎に表したグラフです、右側に赤字で書いてあるのが準拠点数です。

このグラフでわかる事は準拠点数が多いほど復元精度が高い、準拠点重心に近いほど復元精度が高いという当たり前の事が分かるグラフです。

このことからいくつかの組み合わせで準拠点選択が成立していた場合あるいは正規分布になった場合は準拠点数の多い組み合わせを採用すれば安定した精度の高い復元値が得られる理由です。

統計的準拠点選択方法

²正規分布適合度検定・ t 異常値検定 (別ファイル参照)

信頼限界による異常値の除去について

有意水準とは

有意水準の値としては、0.05 (= 5%) を用いるのが一般的ですが厳密さが求められる場合では0.01 (1%) などを用いる場合もあります。

有意であるからといって「偶然ではない」と断定できるわけではなく、「偶然とは考えにくい」という意味に過ぎません。

たとえば有意水準5%で有意という場合には、「実際には偶然に過ぎないのに、誤って『意味がある』と判断している」可能性が多くて5%あるといことです。

座標の位置誤差での ²検定、t検定での準拠点選択の場合は5%を使っています。

信頼限界とは

99%信頼区間とは、母集団全体の値がその区間に存在する確率が99%である区間のことをいい、その区間の下限値を下限信頼限界、上限値を上限信頼限界といいます。測量では3σを限界とすることが多いので座標値の場合、二次元正規分布の3σの確率99%を使うことになると考えます。

Henkan . Xlsでは3σを入力すれば99%、2.45σを入力すれば95%の信頼限界が計算できるように設定してあります。

信頼限界を使った準拠点選択は勧めていません。

²正規分布適合度検定、t異常値検定については「²検定、t検定」のファイルをご覧ください。

ここでは 信頼限界による異常値の除去について簡単に解説します。

始めに有意水準について簡単に説明します。

何事もそうですが100%を保証することは出来ませんのである程度の過失と言いますか危険性を持ってその値を保証することと思えば解りますでしょうか。

統計的な処理では有意水準の値としては、0.05 (= 5%) を用いるのが一般的ですが厳密さが求められる場合では0.01 (1%) などを用いる場合もあります。

有意であるからといって「偶然ではない」と断定できるわけではなく、「偶然とは考えにくい」という意味に過ぎません。

たとえば有意水準5%で有意という場合には、「実際には偶然に過ぎないのに、誤って『意味がある』と判断している」可能性が多くて5%あるといことです。

座標の位置誤差での ²検定、t検定での準拠点選択の場合は5%を使っています。

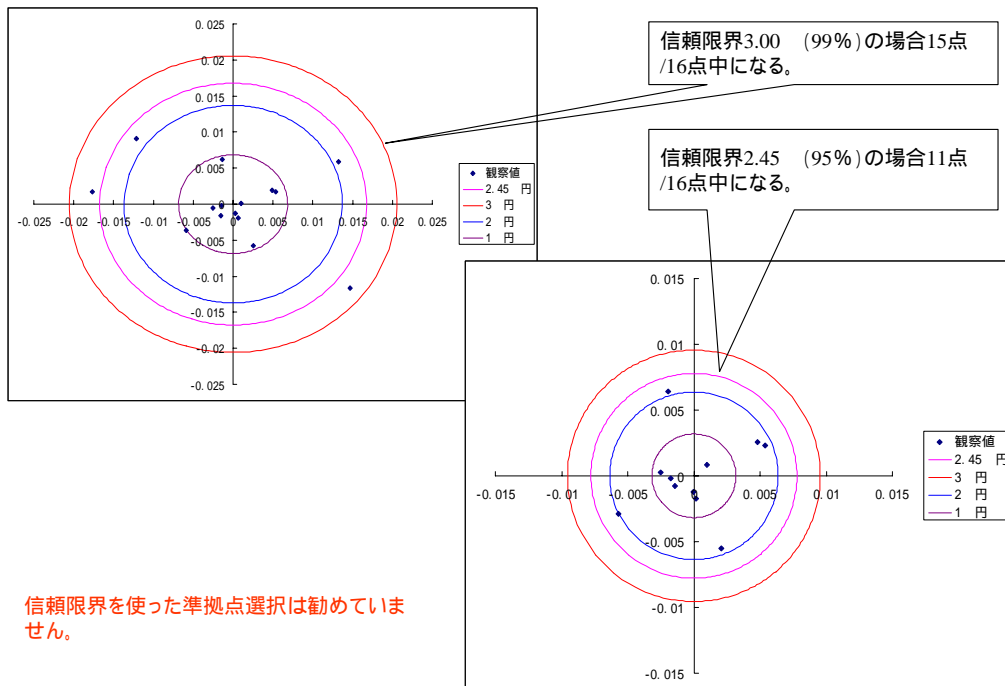
信頼限界とは99%信頼区間とは、母集団全体の値がその区間に存在する確率が99%である区間のことをいい、その区間の下限値を下限信頼限界、上限値を上限信頼限界といいます。

測量では3σを限界とすることが多いので座標値の場合、二次元正規分布の3σの確率99%を使うことになると考えます。

Henkan . Xlsでは3σを入力すれば99%、2.45σを入力すれば95%の信頼限界が計算できるように設定してあります。

が信頼限界を使った準拠点選択は勧めていません。

信頼限界による異常値の除去



この方法が一番鈍くさいといいますが単純な方法です、点数が少ないとき、20点以下の時に使えるときがあります、点数が多いと精度は落ちます。

点数が少ないときは χ^2 検定に近い結果が出ます。

散布図を見れば簡単に判りますので素人には説明しやすいという利点があります、 χ^2 検定、t検定でうまく行かないときに補助的に確認する意味で使っていますのでプログラムにも載せてあります。

図の場合は2.45倍(95%)と3.00倍(99%)で同じデータを点検していますが正直判断が難しいと言うのが本音で、信頼限界を使った準拠点選択はお勧めしません。

これは事例3のデータです。

準拠点の重量

準拠点の重量

多角点、街区点、街区点

点の精度順に準拠点が確保できるのであれば多角点、街区点、筆界点・引証点に沿って準拠点を決めます。

多角点のみ、街区点のみが確保できないが筆界点・引証点に何点が混在している場合で重量をつけたいときの重量の比は(重量は分散の逆数)筆界点1に対して

多角点	4.0	(例 標準偏差 5.0)	多角点	2.0	(例 標準偏差 5.0)
街区点	2.0	(例 標準偏差 7.0)	-----		
		(座標値で確定されている区画整理の場合の街区点)			
筆界点	1.0	(例 標準偏差 10.0)	筆界点	1.0	(例 標準偏差 7.0)

です、精度のチェックをして差があることを確認した上で重量を与えることが必要です。

点の種類による事例は事例3に関連の記載がありますので参考にしてください。

測量の精度は同じ機械を使っていれば単に観測回数の逆数に比例します、この値は分散の逆数と同じになります。

よって多角点は2対回、往復2読定、街区点は1対回、片側2読定、筆界点は半対回、1読定程度と考えれば重量はでます。

最近が機械の性能が高いので4, 2, 1 というよりは2, 1, 1に近いと思います。

準拠点の配置

準拠点の配点バランス

準拠点がある一定の個所に集中していることは適正な状態にありませんので、それは根拠のない重量を与えていることと同じ結果になるからです。

点検方法は図を3～4分割して点の数が二倍以上の差がないことが必要でベクトル図を見て確認出来ます。

ベクトル図を見るポイント

準拠点に偏在が無いが、ベクトル線の方向に偏りが無いかを確認する

誤差の三公理から

- 1．絶対値の等しい正の誤差と負の誤差との起こる度数は相等しい（ベクトル線の方向が360度に渡ってバラバラに向いていること）。
- 2．絶対値の小さな誤差の方が大きい誤差より現れる度数が多い（ベクトル線の長さがこの原則にあっていること）。
- 3．ある程度以上の大きな誤差は実際上起こらない（確率から考えて他と比較して大きなベクトル線があってはならない）。

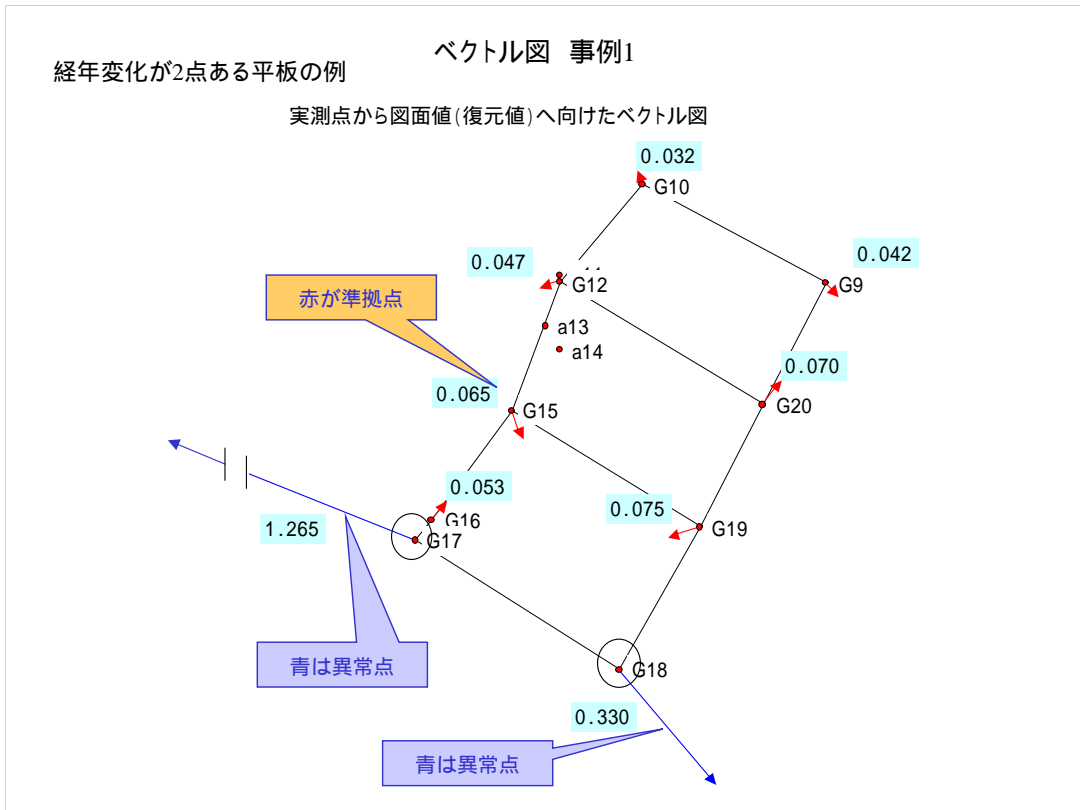
統計的に準拠点を選ぶ方法は「 χ^2 検定、t検定」で解説してありますので参照ください。準拠点は同じ密度で満遍なく配置されていることが理想ですが現実にはそのような配置を期待できません。

が次のことに注意して準拠点の状態を確認する必要があります。

準拠点選択結果を変動ベクトル図によって確認することが重要です。

準拠点に偏在が無いが、ベクトル線の方向に偏りが無いかを確認することがポイントです、あれば図面值か実測値に何らかの問題点が隠されていることとなりますのでその原因を調べて対処しなければなりません。

当然、異常値は除いて考えますが異常点だけでも点の偏在、ベクトルの偏りがあればそこにも違った原因があることとなります、ベクトル図は数値だけの判断で漏れことのある現象を確認するのに非常に便利なアイテムです。

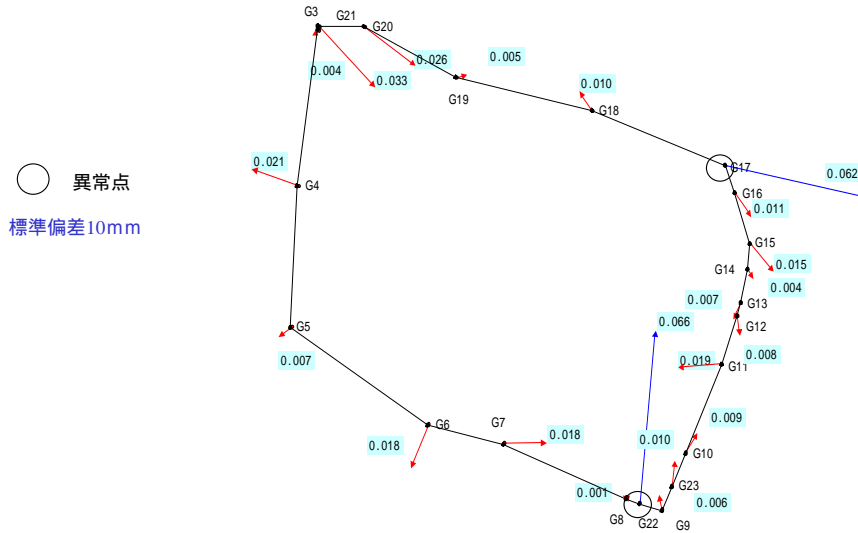


このデータは平板測量の地積測量図から起こしたものです。
 赤の準拠点だけを見れば準拠点に偏在、ベクトル線の方向に偏りは確認出来ませんので
 問題は無いと判断します。

ベクトル図 事例2から

経年変化が2点あるTS測定の例

ベクトル図で準拠点において方向と距離に偏りが無いこと、準拠点が偏って分布していないかを確認します(異常点を除いて)。



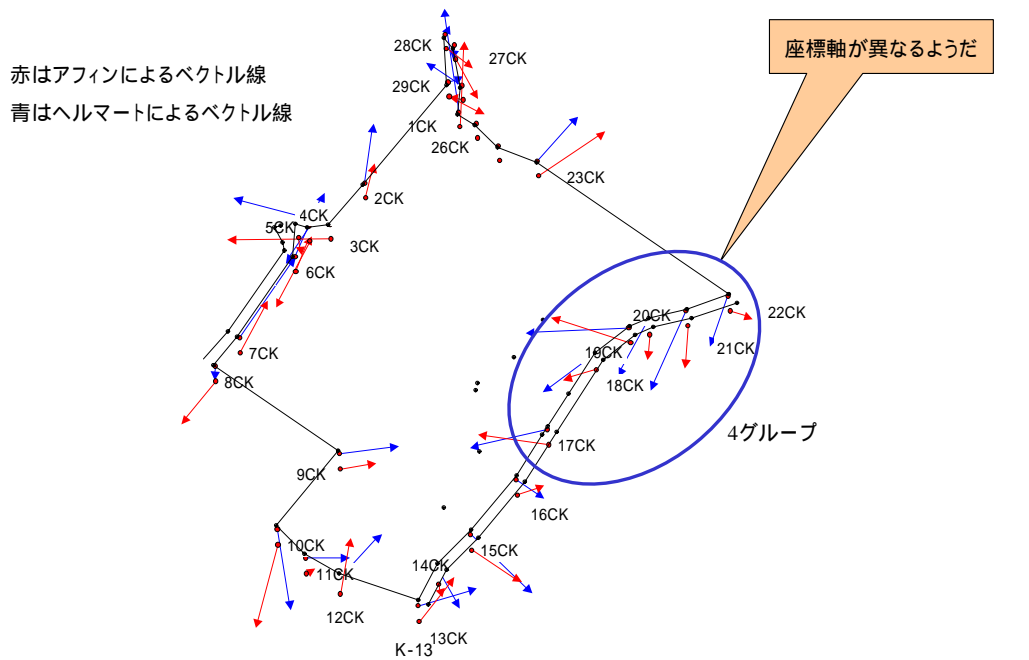
TSの測量結果を10年後に再度測量した結果です、異常点として除いたG17、G22がありません。

ポイントは異常点を除いた点でベクトル線に方向性が無いか、準拠点の配点に偏りが無いか確認します、ここでは問題となるような傾向は見られません。

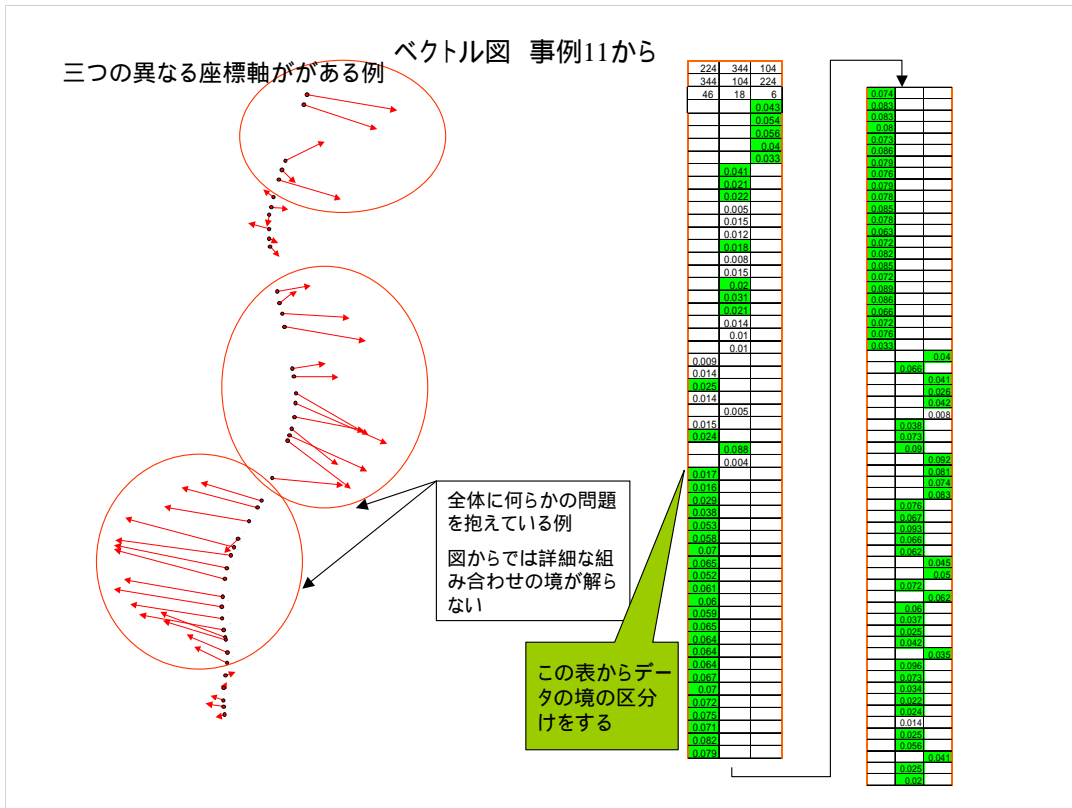
この図では図の右半分に点が集中しているように見えますが左半分に8点 (g 19 ~ g 21、g 3 ~ g 7) と右半分に (g 8 ~ g 18、g 2 2 と g 2 3、G22とG17を除く) 11点の振り分けで問題はありません。

ベクトル図 事例10から

二つの異なる座標軸があるTS測量の例



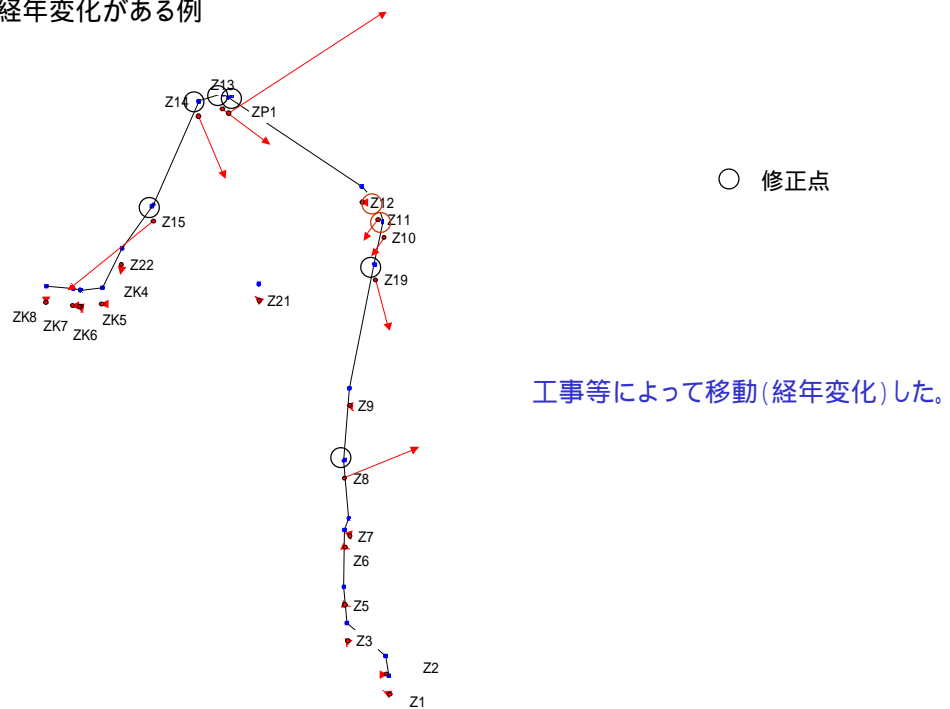
事例10のデータです、青丸の箇所のベクトル線が一定方向を向いている例です。
このあたりが一定の傾向を向いている、つまり座標軸の回転があったと考えられる訳です。
とすればこのまま準拠点として選択するのは適切ではないということです。



事例11のベクトル図です、ベクトル線の向きに偏りがある例で何らかの問題を抱えていることが簡単に分かります。
 この例ではトラバー点がS字型に湾曲している例です、長尺物の道路などに見られる例です。
 このような状態でそのままヘルマート変換とかアフィン変換を掛ける方がいますが注意が必要です。

ベクトル図 事例12から

与点と同じで経年変化がある例



事例12のデータです、座標変換を使わずに同じ与点から多角点を展開して各境界標の移動を確認しているものです。

これは工事によって動いた点を確認しています。

この配点バランスでは座標変換によっても同じ結果が得られますが動いた点が動かない点より多い場合は座標変換によっては確認できません。

統計的な検定が常に正しくないこともあるということを知っておくことが重要です。

異常点が全体の30%を越えると異常点の影響を受けますので層別等の解析が必要です。