

観測データはどれだけの情報量を持っているか? (1): InSAR 及び一般のデータにおける共分散の重要性

How much information does an observed data set include? (1): the importance of covariance for InSAR and other observed data

深畑 幸俊 [1]; 八木 勇治 [2]

Yukitoshi Fukahata[1]; Yuji Yagi[2]

[1] 東大・理・地球惑星; [2] 筑波大学大学院

[1] Dept. Earth and Planet. Science, Univ. Tokyo; [2] Univ. of Tsukuba

ABIC (Akaike's Bayesian Information Criterion, 赤池の情報量規準) を用いたインバージョン解析では、観測データによる情報と先験的情報との重みを ABIC 最小という規準により決める。精度の良いデータが多量にあればデータを良く説明するモデルが選ばれ、逆に誤差が大きかったりデータ数が少なかったりすると先験的情報に従うモデルに近づく。ABIC 導入以前のインバージョン解析では、両者の重みは解析者の主観により選択されていた訳だが、ABIC を用いることにより統計学に基づいて客観的にモデル選択が行えるようになったのが強みである。しかし、その長所を欠点に転じさせないためには、データの持つ情報量の取り扱いについて、これまでとは違ったより深い配慮が必要となる。

まず、InSAR(干渉 SAR) による地震地殻変動データについて考えてみよう。InSAR では植生がないなどの条件が整えば、面的に密にデータを得ることができる。ここで一般にデータのサンプリング間隔が問題となる。例えば、サンプリング間隔を密にすればデータ数が増えるため、見かけ上データの持つ情報量が増える。それを基にインバージョン解析を行うと、当然データを過度に説明するモデルが選択されてしまう。

ところで、InSAR データは主に大気遅延による少なくない誤差を含む。そして、その誤差は空間的に強い相関を持つと考えられる。つまり、個々のデータは完全に独立ではなく、その誤差は無視できない共分散成分を持つ。データ点に近いほど共分散は大きくなり、もしも共分散の値が自身の誤差の大きさと一致してしまえば、そのデータの片方は全く情報を持たないことになる。この共分散成分を取り入れてインバージョン解析を行うことにより、データの持つ情報量を適切に減らすことができると考えられる。

それでももし仮に、大気遅延量が精密に測られるなどして InSAR データが完璧に正確に、つまり誤差無しで地殻変動を表現しているとしたらどうなるであろうか。このような想定は思考実験としては可能である。この時、誤差無しなのでデータの共分散成分はもはや存在せず、データを密に取るほどデータの情報量自体は増えると考えられる。それではこの時、サンプリング間隔を小さく取れば取るほど高精度のインバージョン解析ができるのであろうか。答は否である。そのことを観測方程式を基に説明してみよう。

まず、観測方程式は、

$$d = Ha + e$$

と書ける。d がデータ、a がモデルパラメタ、e が誤差を表すベクトルで、H は適当な理論式 (グリーン関数) に基づいて得られる行列である。問題は、誤差 e の起源で、それにはいくつかのものがある。

例えば、初めに述べた通常の InSAR の観測では、真の地表変位 d_0 に対して、観測で得られる地表変位 d との間には、無視できない誤差 (観測誤差 e_0) がある。これは、

$$e_0 = d_0 - d$$

と表すことができ、上述のように大きな共分散成分も持つ。

次に、理想的状態として観測誤差がない場合 ($d=d_0$) を考える。この時、モデルパラメタ a により真の滑り分布が完全に再現できたとしても ($a=a_0$)、グリーン関数の計算には必ず誤差が含まれる。つまり、真のグリーン関数によって得られる行列 H を H_0 とすると、

$$d_0 = H_0 * a_0$$

$$d_0 = H * a_0 + e_m$$

よって、モデリング誤差 e_m は、

$$e_m = (H_0 - H) * a_0$$

となる。 e_m は、構造モデル等に応じて異なった値となるが、一般に時間的・空間的に何らかの相関を持ったものとなるはずである。要するに、いくら地殻変動が正確に測定できたとしても、我々は完璧な構造モデルを持っていないため、観測方程式に現れる誤差 e は常に共分散成分を持ち、サンプリング間隔をある程度以上小さく取っても意味がない。

地震波データの場合は更に事態が複雑である。地震波解析に用いる H には弾性応答成分に加え、周波数毎に異なる依存性を持つ減衰の効果が含まれているからである。八木・深畑 (本大会発表) が示すように、地震波データでは減衰の効果を考慮することが重要であり、それが高周波サンプリングの限界を規定していると考えられる。いずれにせよ、観測誤差が非常に小さい場合はモデリング誤差の評価が重要であり、モデリング誤差の統計的性質が分かれば、適切なインバージョンを行えると考えられる。