$$\underline{A} = \underline{L}\,\underline{L}^T$$

$\underline{A}$ معلوم و $\underline{L}$ را می‌بایست پیدا کرد :

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ l_{31} & l_{32} & l_{33} & \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \times \begin{bmatrix} l_{11} & l_{12} & l_{13} & l_{14} \\ & l_{22} & l_{23} & l_{24} \\ & & l_{33} & l_{34} \\ & & & l_{44} \end{bmatrix}$$

با انجام ضرب سمت راست تساوی و مساوی قرار دادن عناصر نظیر دو ماتریس داریم :

$$a_{11} = l_{11}^2 \qquad \Rightarrow l_{11} = (a_{11})^{1/2}$$

$$a_{21} = l_{21} l_{11} \qquad \Rightarrow l_{21} = a_{21} / l_{11}$$

$$a_{31} = l_{31} l_{11} \qquad \Rightarrow l_{31} = a_{31} / l_{11}$$

$$a_{41} = l_{41} l_{11} \qquad \Rightarrow l_{41} = a_{41} / l_{11}$$

$$a_{22} = l_{21}^2 + l_{22}^2 \qquad \Rightarrow l_{22} = (a_{22} - l_{21})^{1/2}$$

$$a_{32} = l_{31} l_{21} + l_{32} l_{22} \qquad \Rightarrow l_{32} = (a_{32} - l_{31} l_{21}) / l_{22}$$

$$a_{42} = l_{41} l_{21} + l_{42} l_{22} \qquad \Rightarrow l_{42} = (a_{42} - l_{41} l_{21}) / l_{22}$$

$$a_{33} = l_{31}^2 + l_{32}^2 + l_{33}^2 \qquad \Rightarrow l_{33} = (a_{33} - l_{31}^2 - l_{32}^2)^{1/2}$$

$$a_{43} = l_{41} l_{31} + l_{42} l_{32} + l_{43} l_{33} \qquad \Rightarrow l_{43} = (a_{43} - l_{41} l_{31} - l_{42} l_{32}) / l_{33}$$

$$a_{44} = l_{41}^2 + l_{42}^2 + l_{43}^2 + l_{44}^2 \qquad \Rightarrow l_{44} = (a_{44} - l_{41}^2 - l_{42}^2 - l_{43}^2)^{1/2}$$

مرحلهٔ دوم :

$$\underline{L}\ \underline{O} = \underline{I}$$

$$\begin{bmatrix} l_{11} & & & \\ l_{21} & l_{22} & & \\ l_{31} & l_{32} & l_{33} & \\ l_{41} & l_{42} & l_{43} & l_{44} \end{bmatrix} \times \begin{bmatrix} o_{11} & & & \\ o_{21} & o_{22} & & \\ o_{31} & o_{32} & o_{33} & \\ o_{41} & o_{42} & o_{43} & o_{44} \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

با انجام ضرب سمت چپ تساوی و مساوی قرار دادن عناصر

نظیر داریم :

$$l_{11} o_{11} = 1 \qquad \Longrightarrow o_{11} = 1/l_{11}$$

$$l_{22} o_{22} = 1 \qquad \Longrightarrow o_{22} = 1/l_{22}$$

$$l_{33} o_{33} = 1 \qquad \Longrightarrow o_{33} = 1/l_{33}$$

$$l_{44} o_{44} = 1 \qquad \Longrightarrow o_{44} = 1/l_{44}$$

$$l_{21} o_{11} + l_{22} o_{21} = 0 \qquad \Longrightarrow o_{21} = -o_{22}(l_{21} o_{11})$$

$$l_{31} o_{11} + l_{32} o_{21} + l_{33} o_{31} = 0 \qquad \Longrightarrow o_{31} = -o_{33}(l_{31} o_{11} + l_{32} o_{21})$$

$$l_{41} o_{11} + l_{42} o_{21} + l_{43} o_{31} + l_{44} o_{41} = 0$$

$$\Longrightarrow \qquad o_{41} = -o_{44}(l_{41} o_{11} + l_{42} o_{21} + l_{43} o_{31})$$

$$l_{32} o_{22} + l_{33} o_{32} = 0 \qquad \Longrightarrow o_{32} = -o_{33}(l_{32} o_{22})$$

$$l_{42} o_{22} + l_{43} o_{32} + l_{44} o_{42} = 0 \qquad \Longrightarrow -o_{42} = -o_{44}(l_{42} o_{22} + l_{43} o_{32})$$

$$l_{43} o_{33} + l_{44} o_{43} = 0 \qquad \Longrightarrow o_{43} = -o_{44}(l_{43} o_{33})$$

مرحلهٔ سوم :

$$\underline{B} = \underline{o}^T\ \underline{o}$$

$$\begin{bmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \\ b_{31} & b_{32} & b_{33} & b_{34} \\ b_{41} & b_{42} & b_{43} & b_{44} \end{bmatrix} = \begin{bmatrix} o_{11} & o_{12} & o_{13} & o_{14} \\ & o_{22} & o_{23} & o_{24} \\ & & o_{33} & o_{34} \\ & & & o_{44} \end{bmatrix} \times \begin{bmatrix} o_{11} & & & \\ o_{21} & o_{22} & & \\ o_{31} & o_{32} & o_{33} & \\ o_{41} & o_{42} & o_{43} & o_{44} \end{bmatrix}$$

$$b_{11} = o_{11}^2 + o_{21}^2 + o_{31}^2 + o_{41}^2$$

$$b_{21} = o_{22} o_{21} + o_{32} o_{31} + o_{42} o_{41}$$

$$b_{31} = o_{33} o_{31} + o_{43} o_{41}$$

$$b_{41} = o_{44} o_{41}$$

$$b_{12} = b_{21}$$

$$b_{22} = o_{22}^2 + o_{32}^2 + o_{42}^2$$

$$b_{32} = o_{33} o_{32} + o_{43} o_{42}$$

$$b_{42} = o_{44} o_{42}$$

$$b_{13} = b_{31}$$

$$b_{23} = b_{32}$$

$$b_{33} = o_{33}^2 + o_{43}^2$$

$$b_{43} = o_{44} + o_{43}$$

$$b_{14} = o_{41}$$

$$b_{24} = o_{41}$$

$$b_{34} = o_{43}$$

$$b_{44} = o_{44}^2$$

منابع

1, Ayres, F.(1962). Theory and Problems of Matrices. Schaum's outline Series, McGraw-hill, Toronto.

2, Thompson,E. H. (1969). Introduction to the Algebra of Matrices with some Applicationso University of Toronto Press.

3, Rao, C. R., and Mitra, S. K. (1971). Generalized Inverse of Matrices and Its Applications, John Wiley & Sons , New York.

4, Kurosh, A. (1980). Higher Algebra, Mir Publishers, Moscow.

5, Jaggi, V. P. and Mathur, A. B. (1985). Advanced Engineering Mathematics, Khanna Publishers, Delhi.

فصل دوم

سرشکنی کمترین مربعات

(مطالب مقدماتی و پایه)

۲/۱- مراحل انجام یک پروژهٔ نقشه‌برداری

اکثر فعالیتهائی که نقشه‌بردار در قالب وظایف حرفه‌ای خود انجام می‌دهد نیازمند اندازه‌گیری است . جامعه به صورت سنتی به نقشه‌بردار به دیدهٔ فردی ماهر و متخصص در اندازه‌گیری زمین و عوارض مصنوعی و طبیعی موجود بر آن می‌نگرد . بدین خاطر از او انتظار جمع‌آوری اطلاعات هندسی زمینی را دارد . بالطبع این خواسته نیازمند اندازه‌گیری و به کار گیری آنها در فرمولهای ریاضی است .

اندازه‌گیری به مفهوم واقعی چیزی بیش از آن است که به نظر می‌آید . اگر چه از دیدگاه یک اپراتور ممکن است اندازه‌گیری تماما به عملیات صحرائی محدود شود ، اما یک فرد متخصص لازم است دیدی عمیق‌تر داشته و اندازه‌گیری‌های صحرائی را تنها بخش کوچکی از کار اندازه‌گیری و قدمی در جهت تامین اطلاعات مورد نیاز مراحل بعدی به حساب آورد . به علاوه پیش از شروع به اندازه‌گیری صحرائی لازم است مراحل و چگونگی انجام مشاهدات با توجه به هدف امکانات و دقتهائی که نهایتا مورد نیاز می‌باشد طراحی گردد .

با چنین نگرشی است که می‌توان مراحل عملیات اجرائی یک پروژهٔ نقشه‌برداری را به درستی درک کرد . در نهمین کنفرانس استادان نقشه‌برداری آمریکای شمالی چنین بیان گردید :

"اندازه‌گیری بخشی از مراحل سرشکنی و آنالیز مشاهدات است"

(نگارهٔ ۱) . ونیچک و کراکفسکی در کتاب خود [۲] جزئی از این

مراحل را در قالب عملیات ژئودزی گنجانده اند .

به هر حال مراحل انجام یک پروژهٔ نقشه‌برداری (نقشه‌برداری به مفهوم عام خود که شامل کلیهٔ شاخه‌های آن می‌گردد) را می‌توان به صورت زیر خلاصه کرد .

۱ـ اولین مرحله شناخت کمیتهای مجهول و دقت آنها است .

۲ـ مرحلهٔ بعدی یافتن مدلی است که به کمک آن بتوان به از طریق کمیتهای قابل اندازه‌گیری به کمیتهای مجهول رسید (این مدل را اصطلاحا مدل ریاضی می‌نامیم) . لزوم وجود این مرحله از آنجاست که غالبا امکان اندازه‌گیری مستقیم کمیتهای مجهول وجود ندارد .

۳ـ قبل از شروع عملیات می‌بایست نحوهٔ انجام مشاهدات را طراحی کرد . این عمل را اصطلاحا تجزیه و تحلیل اولیه (Preanalysis) می‌نامیم .

۴ـ پس از انجام طراحی ، مشاهدات بر اساس دستورالعمل تعیین شده در مرحلهٔ طراحی (مرحلهٔ ۳) انجام می‌گیرد .

۵ـ بعد از انجام مشاهدات بلافاصله محاسبات اولیه آغاز می‌گردد . این محاسبات برای دو منظور اساسی صورت می‌گیرند ، کنترل مشاهدات از نظر وجود اشتباه یا خطای سیستماتیک و حذف آن و دیگر کنترل دقت مشاهدات ، اینکه آیا مشاهدات با دقت پیش بینی شده در طرح انجام گرفته اند ؟ یا خیر ؟

۶ـ در این مرحله اطلاعات جمع آوری شده به مدلهای ریاضی معرفی و کمیتهای مجهول محاسبه می‌گردند .

۷ـ پس از محاسبهٔ کمیتهای مجهول نوبت به بررسی و کنترل کیفیت نتایج می‌رسد . برای این منظور آزمونهای آماری متعددی ترتیب

داده خواهد شد .

۸ـ مرحلهء نهائی ارائهء اطلاعات به فرم قابل استفاده به متقاضی است .

همانگونه که در نگارهء ۱ ملاحظه می‌گردد تنها در مرحلهء هفت است که می‌توان مشخص ساخت که آیا نیاز به اندازه‌گیری مجدد و بازگشت به مرحلهء یک وجود دارد ؟ یا خیر ؟

در این مرحله ممکن است لازم شود که حتی در مدل ریاضی و طراحی تجدید نظر کنیم .

۲/۲ـ پردازش اولیهء اطلاعات

از نکات بسیار مهم در به کار گیری روش کمترین مربعات حذف کلیهء خطاهای سیستماتیک و اشتباهات قبل از سرشکنی است. چرا که همانگونه در بخشهای بعدی این فصل خواهید دید جوابی که از طریق این روش برای مجهولات حاصل می‌گردد هنگامی نااریب (Unbiased) خواهد بود که مشاهدات تنها دارای خطای اتفاقی بوده و عاری از هر گونه اشتباه یا خطای سیستماتیک باشند .

به علاوه چون هزینهء انجام مشاهدات بالا است . تکرار مشاهدات بعد از خروج اکیپهای مشاهده از صحرا شدیدا غیر اقتصادی می‌باشد .

بدین خاطر در وهلهء اول روش انجام مشاهدات بایستی به گونه‌ای باشد که به صور مختلف وجود اشتباهات را کنترل کند . به علاوه لازم است در صحرا در حین و بلافاصله بعد از انجام مشاهده

کنترلهای محاسباتی ساده‌ای چون مقایسهٔ نتایج حاصل از تکرار با یکدیگر و برقراری شروط هندسی و ریاضی موجود بین مشاهدات انجام پذیرد .

به این محاسبات و کنترلها پردازش اولیه (یا Pre-processing) اطلاق می‌گردد . مجموعهٔ اعمالی که تحت این عنوان صورت می‌گیرند محدودهٔ وسیعی دارند . بطوری که امروزه حتی امکان سرشکنی مقدماتی مشاهدات با کامپیوترهای قابل حمل در صحرا وجود دارد. به این ترتیب می‌توان از صحت و دقت نتایج نیز مطمئن شد . به علاوه در صورت نیاز به امکانات کامپیوتری بیشتر ارسال این اطلاعات از طریق خط تلفن به دفتر مرکزی و انجام محاسبات نهایی پیش از ترک اکیپهای مشاهده از منطقهٔ عملیات میسر است .

۲/۳- علت نیاز به سرشکنی (سرشکنی مدل پارامتریک خطی)

به سه دلیل در اندازه‌گیری به حداقل مشاهدات لازم اکتفا نمی گردد :

- ایجاد امکان کشف اشتباه (تنها راه کشف اشتباه تکرار است).
- رسیدن به برآورد نزدیک به واقع از پارامترهای مجهول (هر مشاهده به منزله یک عضو از جوامعهٔ با بی‌نهایت عضو است ، بنابر این هر چه تعداد اعضاء نمونه بیشتر باشد ، امکان نتیجه‌گیری بر

مبنای اعضاء نادر جامعه کمتر خواهد شد).

- ایجاد امکان برآورد دقت اندازه‌گیری و نتایج حاصل.

فرض کنید رابطه بین مشاهدات ومجهولات بصورت زیر باشد:

$$\underline{l}=\underline{A}\underline{x} \qquad (۱)$$

$\underline{x}$ بردار مجهولات، $\underline{l}$ بردار مشاهدات و $\underline{A}$ ماتریس ضرائب مجهولات است. به معادلاتی به شکل فوق، معادلات پارامتریک یا معادلات مشاهدات (Observations Equations یا Parametric Equations) اطلاق می‌گردد. اگر تعداد مجهولات این مدل را u در نظر بگیریم تنها u مشاهده برای حل دستگاه و یافتن مجهولات از طریق معادلات (۱) کافی خواهد بود. اما به دلایل ذکر شده اکتفای به حداقل مشاهدات لازم، صحیح نیست. لذا این تعداد مشاهدات (یا معادلات) (n) را بیش از تعداد مجهولات افزایش می‌دهیم. اما از آنجایی که مشاهدات بالطبع آلوده به خطاهای اتفاقی‌اند (البته پس از حذف اشتباهات و خطاهای سیستماتیک) معادلات ناسازگار شده و لذا جواب نخواهد داشت. چون:

$$\underline{D}=[\underset{(n*u)}{\underline{A}} \quad \underset{(n*1)}{\underline{l}}] \qquad [n*(u+1)]$$ ماتریس بسط داده شده

$rank(D)=u+1$ (بشرط آنکه ستونهای $\underline{A}$ صفر و یا وابسته نباشد)

$rank(A)=u$ (بشرط وابسته یا صفر نبودن ستونهای $\underline{A}$)

$$\Rightarrow \quad rank(\underline{A}) \neq rank(\underline{D})$$

و بنابه آنچه که در فصل (۱) در مورد تعداد جوابهای یک دستگاه ذکر شد، معادلات (۱) ناسازگار بوده و فاقد جوابند. تنها راه خروج از وضعیت حاصل و یافتن نوعی جواب برای (۱)

پذیرش آلوده به خطا بودن مشاهدات و اقدام به تصحیح آنهاست، یعنی افزودن بردار مجهول $\underline{v}$ به معادلات (۱):

$$\underline{v}+\underline{l}=\underline{A}\underline{x} \qquad (۲)$$

و این آغاز مبحث سرشکنی (یا Adjustment) است. به بردار $\underline{v}$، بردار تصحیحات یا باقیمانده‌ها (Residuals) گفته می‌شود. دستگاه معادلات (۲) که به آن دستگاه معادلات تغییر شکل یافته (Reformulated Equations) اطلاق می‌گردد دارای بی‌نهایت جواب است. چون با توجه به مجهولات جدید، معادله (۲) را می‌توان به شکل زیر نوشت.

$$\underline{l}=[\underline{A}\ -\underline{I}]\begin{bmatrix}\underline{x}\\ \underline{v}\end{bmatrix}=\underline{N}\begin{bmatrix}\underline{x}\\ \underline{v}\end{bmatrix}$$

حال برای تعیین تعداد جواب، $\mathrm{rank}(\underline{D})$ (ماتریس بسط داده شده) را با $\mathrm{rank}(\underline{N})$ (ماتریس ضرائب جدید) مقایسه می‌کنیم.

$\mathrm{rank}(\underline{N})=\mathrm{rank}[\underline{A}\ -\underline{I}]=n$ (چون $\det(-\underline{I})\neq 0$ است)

$\mathrm{rank}(\underline{D})=\mathrm{rank}[\underline{A}\ -\underline{I}\ \underline{l}]=n$ (" " ")

اما تعداد مجهولات u+n است (بردار مجهولات شامل $\underline{x}$ و $\underline{v}$ است). بنابراین دستگاه (۲) بی‌نهایت جواب خواهد دارد. نکته قابل توجه آنستکه $\underline{l}$ بردار مشاهدات بوده و تصحیح آن به هر مقدار جایز نیست. اگر برای این تصحیحات شرط

$$\underline{v}^T\underline{v}=\sum_{i=1}^{n}\underline{v}_i^2=\min \qquad (۳)$$

را قائل شویم اولا تنها یک $\underline{x}$ دارای این خاصیتی است که $\underline{v}$ های حاصل برای آن در شرط (۳) صدق کنند و بعلاوه به این ترتیب مشاهدات را با مقادیری تصحیح کرده‌ایم که مجموع مربعات آنها

حداقل است . به این جواب ، جواب کمترین مربعات (Least Squars) اطلاق می‌شود . توجه داشته باشید که میتوان برای باقیمانده‌ها شروطی غیر از شرط (۳) را نیز در نظر گرفت ، اما در آنصورت به جواب حاصل جواب کمترین مربعات گفته نخواهد شد . یکی از شروطی که می‌تواند جایگزین شرط کمترین مربعات گردد شرط ذیل است.

$$\sum_{i=1}^{n}|v_i|=\min$$

به جواب حاصل از این شرط جواب پایدار (Robust) گفته می‌شود چون کمتر تحت تاثیر خطای سیستماتیک و اشتباهات است. برای بدست آوردن $\underline{x}$ هایی که شرط (۳) را برآورده می سازند ، از (۳) نسبت به بردار $\underline{x}$ مشتق گرفته مساوی صفر قرار می‌دهیم .

$$\frac{d(\underline{v}^T\underline{v})}{d\underline{x}}=\frac{d(\underline{A}\underline{x}-\underline{l})^T(\underline{A}\underline{x}-\underline{l})}{d\underline{x}}$$

$$=\left(\frac{d(\underline{A}\underline{x}-\underline{l})^T(\underline{A}\underline{x}-\underline{l})}{d(\underline{A}\underline{x}-\underline{l})}\right)\left(\frac{d(\underline{A}\underline{x}-\underline{l})}{d\underline{x}}\right)$$

$$=2(\underline{A}\underline{x}-\underline{l})^T\underline{A}=0$$

در صورتیکه ستونهای ماتریس $\underline{A}$ صفر یا وابسته نباشند آنگاه $(\underline{A}^T\underline{A})$ غیر مجرد بوده و داریم :

$$\hat{\underline{x}}=(\underline{A}^T\underline{A})^{-1}\underline{A}^T\underline{l} \qquad (۴)$$

گفته شد که اگر ماتریس $\underline{A}$ دارای ستونهای وابسته و صفر نباشد $(\underline{A}^T\underline{A})$ غیر مجرد است . برای اثبات این مطلب $\underline{A}$ را بصورت زیر می‌نویسیم :

$$\underline{A}=\begin{bmatrix} a_{11} & \cdots & a_{1u} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = [\underline{a}_1 \cdots \underline{a}_u]$$

ودر نتیجه $\underline{A}^T\underline{A}$ بصورت زیر قابل محاسبه است:

$$\underline{A}^T\underline{A} = \begin{bmatrix} \underline{a}_1^T\underline{a}_1 & \underline{a}_1^T\underline{a}_2 & \cdots & \underline{a}_1^T\underline{a}_u \\ \vdots & \vdots & & \vdots \\ \underline{a}_u^T\underline{a}_1 & \underline{a}_u^T\underline{a}_2 & \cdots & \underline{a}_u^T\underline{a}_u \end{bmatrix}$$

که دترمینان آن تنها هنگامی مساوی صفر است که یکی از بردارهای $\underline{a}_1,\cdots,\underline{a}_u$ مساوی صفر و یا وابسته باشند.

تمرین :

نشان دهید در صورتیکه دو ستون $\underline{A}$ وابسته باشند (وابستگی‌خطی داشته باشند) $\det(\underline{A}^T\underline{A})=0$ است.

برای بررسی آنکه جواب حاصل از (۴) ماکزیمم کننده $(\underline{v}^T\underline{v})$ است، یا مینیمم کننده آن، علامت مشتق دوم آنرا به ازای $\underline{x}$ مورد بررسی قرار می‌دهیم.

$$\phi = \underline{v}^T\underline{v} = (\underline{A}\,\underline{x}-\underline{l})^T(\underline{A}\,\underline{x}-\underline{l})$$

$$\frac{d\phi}{d\underline{x}} = \frac{d\phi}{d(\underline{A}\underline{x}-\underline{l})} = \frac{d(\underline{A}\,\underline{x}-\underline{l})}{d\underline{x}} = 2(\underline{A}\,\underline{x}-\underline{l})^T\underline{A}$$

$$=2(\underline{x}^T\underline{A}^T-\underline{l}^T)\underline{A}=2(\underline{x}^T\underline{A}^T\underline{A}-\underline{l}^T\underline{A})$$

$$\frac{d^2\phi}{d\underline{x}^2} = 2\underline{A}^T\underline{A}$$

حال بایستی علامت $\underline{A}^T\underline{A}$ را تعیین کرد. $\underline{A}^T\underline{A}$ یک ماتریس مربع $u \times u$ است، و تعیین علامت آن از طریق تعیین علامت مقادیر ویژه آن میسر است. بنابر فرضی که برای وجود جواب (۱) کرده ایم:

$$\underline{A}=[\ \underline{a}_1\ \underline{a}_2 \cdots\cdots \underline{a}_u\ ]$$

بطوریکه $\underline{a}_1, \cdots\cdots, \underline{a}_u$ مستقل اند.

$$\underline{A}^T\underline{A} = \begin{bmatrix} \underline{a}_1^T \\ \underline{a}_2^T \\ \vdots \\ \underline{a}_u^T \end{bmatrix} [\underline{a}_1\ \underline{a}_2 \cdots\cdots \underline{a}_u] = \begin{bmatrix} \underline{a}_1^T\underline{a}_1 & \cdots\cdots\cdots\cdots & \underline{a}_1^T\underline{a}_u \\ \vdots & & \\ \vdots & & \\ \underline{a}_u^T\underline{a}_1 & \cdots\cdots\cdots\cdots & \underline{a}_u^T\underline{a}_u \end{bmatrix} \qquad (۵/۲)$$

ماتریس (۵) کلیه مقادیر ویژه اش مثبت اند، چون اولا بخاطر مستقل بودن $\underline{a}_1, \cdots\cdots, \underline{a}_u$ غیر مجرد بوده و بعلاوه مقادیر روی قطر اصلی آن مربع کامل اند. بنابر این $\underline{A}^T\underline{A}$ یک ماتریس مثبت بوده و جواب حاصل از (۴) مینیمم کننده ء شرط (۳) است. به ماتریسی که تمام مقادیر ویژه آن مثبت باشد، اکیدا مثبت (Positive Definite) اطلاق می‌گردد.

ماتریس (۵،۲) یک تنسور متریک (Metrix Tensor) برای فضای مجهولات است، چون مستقل بودن $\underline{a}_1, \cdots, \underline{a}_u$ بیانگر این است که $\underline{a}_1, \cdots, \underline{a}_u$ یک مجموعه بردار پایه (Basis) برای فضای مجهولات می‌باشد و بعلاوه عناصر روی قطر اصلی $\underline{A}^T\underline{A}$ ضرب داخلی بردارهای پایه در یکدیگر و عناصر خارج قطر اصلی حاصلضرب بردارهای پایه در یکدیگرند.

۲/۴- سرشکنی همراه با وزن

جوابی که برای مجهولات بدست آوردیم جوابی برای حالتی خاص بود. حالتی که در آن کلیه مشاهدات ارزش و اعتبار یکسانی داشتند و درنتیجه ملاک تعیین جواب شرط $\underline{v}^T\underline{v}=Min$ بود. اما اگر مشاهدات دارای ارزشهای متفاوت باشند. این ارزشها را می‌توان بر حسب اعدادی که نشان دهندهٔ دقت نسبی مشاهدات است در داخل ماتریس قطری $\underline{P}$ قرارداد.

$$\underline{P} = \begin{bmatrix} p_1 & 0 & \cdots\cdots & 0 \\ 0 & p_2 & \cdots\cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots\cdots & p_n \end{bmatrix}$$

$p_1, \ldots\ldots, p_n$ دقت نسبی مشاهداتند و $\underline{P}$ را ماتریس وزن می‌نامیم. حال برای آنکه دقت و ارزش مشاهدات را در تعیین جواب ملحوظ داریم، شرط ذیل را به مدل پارامتریک تغییر یافته $(\underline{v}+\underline{l}=\underline{A}\ \underline{x})$ می‌افزائیم:

$$\underline{v}^T\underline{P}\ \underline{v}=Min$$

به این ترتیب مشاهدهٔ با ارزش یا وزن بیشتر، تصحیح کمتری خواهد پذیرفت.

$$\phi=\underline{v}^T\underline{P}\ \underline{v}=(\underline{A}\ \underline{x}-\underline{l})^T\underline{P}(\underline{A}\ \underline{x}-\underline{l}) \qquad (۶)$$

برای رسیدن به جواب مینیمم برای ϕ از (۶) نسبت به $\underline{x}$ مشتق

گرفته مساوی صفر قرار می‌دهیم :

$$\frac{d\phi}{d\underline{x}} = \frac{d\phi}{d(\underline{A}\,\underline{x}-\underline{l})} \times \frac{d(\underline{A}\,\underline{x}-\underline{l})}{d\underline{x}} = (\underline{A}\,\underline{x}-\underline{l})^T\,\underline{P}\,\underline{A} = \underline{0}$$

$$(\underline{A}^T\,\underline{P}\,\underline{A})\,\underline{x} = \underline{A}^T\,\underline{P}\,\underline{l}$$

و در صورتیکه $\underline{P}$ و $\underline{A}$ کمبود مرتبه نداشته باشند :

$$\hat{\underline{x}} = (\underline{A}^T\,\underline{P}\,\underline{A})^{-1}\,\underline{A}^T\underline{P}\,\underline{l} \qquad (۷)$$

کمبود مرتبه نداشتن $\underline{A}$ مستلزم مستقل بودن ستونهای آن و در مورد $\underline{P}$ عدم وجود عناصر صفر روی قطر اصلی است (یعنی $\underline{P}$ را بایستی بصورت ماتریسی قطری با عناصر قطر مخالف صفر انتخاب کنیم).

برای بررسی مینیمم کننده بودن $\hat{\underline{x}}$ از (۶) نسبت به $\underline{x}$ دو بار مشتق میگیریم :

$$\frac{d^2\phi}{d\underline{x}^2} = 2\underline{A}^T\underline{P}\,\underline{A} \qquad (۸)$$

برای آنکه مقادیر ویژه (۸) همواره مثبت باشند می‌بایست $\underline{A}$ کمبود مرتبه نداشته و بعلاوه عناصر قطری $\underline{P}$ مثبت باشند. بنابر این وزنهائی که برای مشاهدات انتخاب می‌کنیم بایستی مثبت و مخالف صفر باشند.

بعد از برآورد $\underline{x}$ از رابطه (۷) امکان برآورد باقیمانده‌ها نیز وجود خواهد داشت:

$$\underline{v} = \underline{A}\,\underline{x}-\underline{l} \qquad \hat{\underline{v}} = \underline{A}\,\hat{\underline{x}}-\underline{l}$$

$$= \underline{A}(\underline{A}^T\underline{P}\,\underline{A})^{-1}\underline{A}^T\underline{P}\,\underline{l}-\underline{l} \qquad (۹)$$

بر اساس آنچه که تاکنون اثبات گردید،

$$\hat{\underline{x}} = (\underline{A}^T \underline{P}\, \underline{A})^{-1} \underline{A}^T \underline{P}\, \underline{l}$$

برآورد کمترین مربعات مدل

$$\underline{v}+\underline{l}=\underline{A}\, \underline{x}$$

است. حال میخواهیم ثابت کنیم که:

$$E(\hat{\underline{x}})=\underline{x} \qquad (۱۰)$$

یعنی $\hat{\underline{x}}$ یک برآورد نااریب (Unbiased) از $\underline{x}$، مجهول معادله (۲) است. برای اثبات این مطلب از (۷) امید ریاضی میگیریم:

$$E(\hat{\underline{x}})=E[(\underline{A}^T \underline{P}\, \underline{A})^{-1} \underline{A}^T \underline{P}\, \underline{l}]$$
$$=(\underline{A}^T \underline{P}\, \underline{A})^{-1} \underline{A}^T \underline{P}\, E(\underline{l}) \qquad (۱۱)$$

از طرفی با توجه به (۲):

$$\underline{l}=\underline{A}\, \underline{x}-\underline{v} \qquad (۱۲)$$

بنابراین اگر نتیجه (۱۲) را در (۱۱) قرار دهیم:

$$E(\hat{\underline{x}})=(\underline{A}^T \underline{P}\, \underline{A})^{-1} \underline{A}^T \underline{P}\, E(\underline{A}\, \underline{x}-\underline{v})$$
$$=(\underline{A}^T \underline{P}\, \underline{A})^{-1} \underline{A}^T \underline{P}\, \underline{A}\, \underline{x}-E(\underline{v}) \qquad (۱۳)$$
$$=\underline{x}-E(\underline{v}) \qquad (۱۴)$$

اگر مشاهدات عاری از اشتباه و خطای سیستماتیک باشند، (تنها تحت تاثیر خطاهای اتفاقی باشند) در مدل (۲) $\underline{v}$ خطای

اتفاقی مشاهدات بوده لذا $E(\underline{v})=\underline{0}$ خواهد شد و از (۱۴) نتیجه می‌گیریم که :

$E(\hat{\underline{x}})=\underline{x}$ (در صورت عدم وجود اشتباه و خطای سیستماتیک)

که باعث تکمیل اثبات می‌گردد .

۲/۶- ماتریس کوواریانس باقیمانده‌ها

طبق تعریف ماتریس کوواریانس، در مورد باقیمانده‌ها داریم:

$$\underline{C}_v = E[(\underline{v} - E(\underline{v}))(\underline{v} - E(\underline{v}))^T] \qquad (۱۵)$$

در صورتیکه مشاهدات تنها دارای خطای اتفاقی باشند، $E(\underline{v}) = 0$ بوده و در نتیجه:

$$\underline{C}_v = E(\underline{v}\,\underline{v}^T) \qquad (۱۶)$$

اما با توجه به مدل (۲)،

$$\underline{v} = \underline{A}\,\underline{x} - \underline{l} \qquad (۱۷)$$

با اعمال اپراتور E (امید ریاضی) به طرفین تساوی (۱۷) داریم:

$$E(\underline{l}) = \underline{A}\,\underline{x} \qquad (۱۸)$$

حال اگر در (۱۷) بجای $\underline{A}\,\underline{x}$ مقدار مساوی آنرا از (۱۸) قرار دهیم،

$$\underline{v} = E(\underline{l}) - \underline{l} \qquad (۱۹)$$

و نهایتا،

$$\underline{C}_v = E[(E(\underline{l}) - \underline{l})(E(\underline{l}) - \underline{l})^T] = \underline{C}_l \qquad (۲۰)$$

یعنی ماتریس کوواریانس مشاهدات با ماتریس کوواریانس باقیمانده‌ها برابر است.

۲/۷- جواب حاصل از روش کمترین مربعات
و
برآورد ماکزیمم احتمال باقیمانده‌ها

اگر در مدل پارامتریک (۲) مشاهدات دارای تابع توزیع نرمال باشند، $\underline{v}$ هم که تابعی خطی از $\underline{l}$ است دارای تابع توزیع نرمال خواهد بود. پس در این حالت تابع توزیع چگالی باقیماندها به صورت ذیل خواهد بود.

$$f(\underline{v}) = \underbrace{\frac{(\det(\underline{C}_{l}^{-1}))^{1/2}}{(2\pi)^{n/2}}}_{C} \exp\left[-1/2(\underline{v}-E(\underline{v}))^{T}\underline{C}_{v}^{-1}(\underline{v}-E(\underline{v}))\right]$$

اگر مشاهدات فاقد خطای سیستماتیک باشند $(E(\underline{v})=\underline{0})$:

$$f(v) = C \exp\left[-1/2\ \underline{v}^{T}\underline{C}_{v}^{-1}\underline{v}\right]$$

بنابر این جوابی که مینیمم کننده‌ء $(\underline{v}^{T}\underline{C}_{v}^{-1}\underline{v})$ است ماکزیمم کننده احتمال $\underline{v}$ ها است. یعنی اگر در مدل پارامتریک $(\underline{v} + \underline{l} = \underline{A}\ \underline{X})$ بخواهیم x ای را بدست آوریم که برای آن v احتمال ماکزیمم داشته باشد باز هم به همان جواب کمترین مربعات خواهیم رسید. البته به شرط آنکه ماتریس وزن را عکس ماتریس کوواریانس مشاهدات، و یا متناسب با عکس آن انتخاب کرده باشیم.

$$\underline{P}=\underline{C}_{l}^{-1} \qquad \text{یا} \qquad \underline{P} \propto \underline{C}_{l}^{-1}$$

متاسفانه عملا ندرتا پیش از سرشکنی ماتریس کوواریانس مشاهدات قبل از سرشکنی مطور کامل در اختیار است. و در واقع یکی از اهداف سرشکنی یافتن برآورده‌ای برای ماتریس کوواریانس

مشاهدات یا بطور کلی دقت مشاهدات می‌باشد.

اما همواره قبل از سرشکنی تشکیل ماتریسی که دقت نسبی مشاهدات را نشان دهد میسر است. به این ماتریس که با ماتریس کوواریانس واقعی مشاهدات تنها در یک ضریب (یا مقیاس) متفاوت است، ماتریس کوواریانس نسبی ($\underline{Q}$) می‌گوئیم.

$$\underline{C}_{-1} = \sigma_0^2 \, \underline{Q} \qquad (۲۱)$$

بنابر این حداکثر می‌توان وزن مشاهدات را عکس ماتریس کوواریانس نسبی انتخاب کرد.

$$\underline{P} = \underline{Q}^{-1} \qquad (۲۲)$$

واضح است که این انتخاب بر روی نتیجهٔ اخیر اثری می‌گذارد. اما ببینیم تاثیر آن بر جواب کمترین مربعات چگونه است.

قبلا اثبات کردیم که:

$$\hat{\underline{x}} = (\underline{A}^T \underline{P} \underline{A})^{-1} \underline{A}^T \underline{P} \underline{1}$$

$$= (\underline{A}^T \underline{Q}^{-1} \underline{A})^{-1} \underline{A}^T \underline{Q}^{-1} \underline{1}$$

$$= (\underline{A}^T \sigma_0^2 \underline{C}_{-1}^{-1} \underline{A})^{-1} \underline{A} \; \sigma_0^2 \underline{C}_{-1}^{-1} \underline{1}$$

$$= (\underline{A}^T \underline{C}_{-1}^{-1} \underline{A})^{-1} \underline{A}^T \underline{C}_{-1}^{-1} \underline{1}$$

یعنی نتیجهٔ حاصل از سرشکنی همان خواهد است که از طریق انتخاب ماتریس وزن به صورت عکس ماتریس کوواریانس مشاهدات (صورت دستیابی به آن) حاصل می‌گردد. بنابراین انتخاب وزن بصورت $\underline{P} = \underline{C}_{-1}^{-1}$ یا $\underline{P} = \underline{Q}^{-1}$ منجر به نتیجهٔ یکسان برای مجهولات خواهد شد.

۲/۸- جواب حاصل از روش کمترین مربعات
و
برآورد کننده مینیمم واریانس

فرض میکنیم :

$$\underline{x} = \underline{B}\,\underline{l} \qquad (۲۳)$$

یک بر آورد کننده نااریب برای $\underline{x}$ باشد .

در مدل $(\underline{v} + \underline{l} = \underline{A}\,\underline{x})$، $\hat{x}$ یک بردار بوده و بنابر این مفهوم مینیمم واریانس برای آن معادل تریس (Trace) مینیمم برای $\underline{C}_{\hat{x}}$ است. پس هدف ما این است که برآوردکنندهء نااریبی برای $\hat{\underline{x}}$ بدست آوریم که در مورد آن $\text{trace}\,(\underline{C}_{\hat{x}}) = \min$ باشد .

خاصیت نااریبی حکم میکند که :

$$E\,(\hat{\underline{x}}) = \underline{x} \qquad (۲۴)$$

با اعمال اپراتور $\underline{E}$ بر (۲۳) داریم :

$$E\,(\hat{\underline{x}}) = E\,(\underline{B}\,\underline{l}) = \underline{B}\,E\,(\underline{A}\,\underline{x} - \underline{v}) = \underline{B}\,\underline{A}\,\underline{x}$$

بنابراین فرض نااریبی در صورتی برقرار خواهد بود که $\underline{B}\,\underline{A} = \underline{I}$ باشد .

حال قانون انتشار خطاها را به رابطهء (۲۳) اعمال میکنیم .

$$\underline{C}_{\hat{x}} = \underline{B}\,\underline{C}_{l}\,\underline{B}^{T}$$

اکنون $\underline{B}$ را بگونه ای مییابیم که :

$$\text{Trace}\;(\underline{B}\,\underline{C}_{l}\,\underline{B}^{T})$$

مینیمم شده و در عین حال شرط نااریبی $(\underline{B}\,\underline{A} - \underline{I} = \underline{0})$ برقرار باشد .

برای این منظور از روش لاگرانژ استفاده می‌کنیم :

$$\phi = \underline{B}\,\underline{C}_1\,\underline{B}^T + 2\,(\underline{B}\,\underline{A} - \underline{I})\,\underline{k} = \underline{0}$$

$\underline{k}$ ماتریس با عناصر مجهول است که به آن ماتریس ضرایب لاگرانژ (Lagrange multipliers) گفته میشود .

بر اساس قواعد مشتق از Trace ،که در فصل (۱) نیز مرور گردید ، داریم :

$$Tr\,(\phi) = Tr\,(\underline{B}\,\underline{C}_1\,\underline{B}^T) + 2\,Tr\,(\underline{B}\,\underline{A}\,\underline{k}) - 2\,Tr\,(\underline{k})$$

$$\frac{d\,Tr\,(\underline{B}\,\underline{C}_1\underline{B}^T)}{d\,\underline{B}} = \underline{C}_1(\underline{B}_1^T + \underline{B}_1^T) = 2\,\underline{C}_1\underline{B}^T$$

$$\frac{d\,Tr\,(\underline{B}\,\underline{A}\,\underline{k})}{d\,\underline{B}} = \underline{A}\,\underline{K}$$

$$\frac{d\,Tr\,(\underline{k})}{d\,\underline{B}} = \underline{0}$$

$$\frac{d\,Tr(\phi)}{d\,\underline{B}} = 2\,\underline{C}_1\underline{B}^T + 2\,\underline{A}\,\underline{K} = \underline{0}$$

$$\Rightarrow \underline{B} = -\,\underline{k}^T\,\underline{A}^T\,\underline{C}_1^{-1}$$

چون $\underline{B}\,\underline{A} = \underline{I}$ است پس طرفین تساوی فوق را از راست در $\underline{A}$ ضرب کرده مساوی $\underline{I}$ قرار می‌دهیم :

$$\underline{k}^T = -(\underline{A}^T\underline{C}_1^{-1}\underline{A})^{-1} \qquad \underline{B} = (\underline{A}^T\underline{C}_1^{-1}\underline{A})^{-1}\,\underline{A}^T\underline{C}_1^{-1}$$

و نهایتا :

$$\hat{\underline{x}} = \underline{B}\,\underline{l} = (\underline{A}^T \underline{C}_l^{-1} \underline{A})^{-1} \underline{A}^T \underline{C}_l^{-1} \underline{l}$$

با مقایسه این جواب با جواب حاصل از اعمال شرط کمترین مربعات به این نتیجه می‌رسیم که در صورتیکه $\underline{P} = \underline{C}_l^{-1}$ انتخاب شود جواب حاصل از روش کمترین مربعات دارای کوچکترین واریانس خواهد بود . توجه داشته باشید که در صورت انتخاب وزن بصورت $\underline{P} = \underline{Q}^{-1}$ همچنان نتیجهٔ حاصل صحیح خواهد بود .

۲/۹- جواب کمترین مربعات مدل پارامتریک و شبه معکوس

در مدل پارامتریک ($\underline{l} = \underline{A}\,\underline{x}$) ، $\hat{\underline{x}}$ حاصل از (۷) جوابی منحصر به فرد برای x است. از طرفی برای این مدل جواب ذیل نیز منحصر به فرد است.

$$\hat{\underline{x}} = \underline{A}^+ \underline{l} \qquad (۲۵)$$

حال این سوال به ذهن متبادل می‌گردد که آیا ارتباطی بین ضریب $\underline{l}$ در رابطهٔ (۷) شبه معکوس A نیست؟

برای پاسخ به این سوال چهار خاصیت شبه معکوس را برای ضریب $\underline{l}$ در رابطهٔ (۷) یعنی :

$$\underline{B} = (\underline{A}^T \underline{P}\,\underline{A})^{-1} \underline{A}^T \underline{P} \qquad (۲۶)$$

امتحان می‌کنیم . شروط (i) ، (ii) و (iv) شبه معکوس (رجوع کنید به فصل (۱)) به وضوح برقرارند ، تنها در مورد شرط (iii) داریم :

$$(\underline{A}\,(\underline{A}^T \underline{P}\,\underline{A})^{-1} \underline{A}^T \underline{P})^T = \underline{P}\,\underline{A}\,(\underline{A}^T \underline{P}\,\underline{A})^{-1} \underline{A}^T \qquad (۲۷)$$

که تنها هنگامی برقرار خواهد بود که $\underline{P} = \underline{I}$ باشد . یعنی اگر $\underline{P}=\underline{I}$ باشد جواب حاصل از روش کمترین مربعات و جواب حاصل از (۲۵) یکی خواهند بود .

این نتیجه دستورالعملی برای تعیین شبه معکوس ماتریسهای مستطیلی با تعداد سطرهای بیش از ستون و فاقد کمبود مرتبه در اختیار ما میگذارد .

$$\underline{A}^{+} = (\underline{A}^{T}\underline{A})^{-1}\underline{A}^{T} \qquad (۲۸)$$

۲/۱۰- ماتریسهای کوواریانس کمیتهای حاصل از سرشکنی مدل پارامتریک خطی

همانگونه که قبلا گفته شد یکی از اهداف تکرار در اندازه‌گیری برآورد دقت (precision) مشاهدات و مجهولات است. بدین لحاظ در این بخش می‌خواهیم برآوردکننده‌هائی برای دقت کمیتهای حاصل از سرشکنی کمترین مربعات بدست آوریم. به این ترتیب انتقال آماری از فضای مشاهدات به مجهولات تکمیل می‌گردد (اطلاعات آماری فضای مشاهدات در $\underline{C}_l$, $\underline{l}$ نهفته می‌باشد).

قابل ذکر است که، بدست آوردن ماتریس کوواریانس کمیتهای حاصل از سرشکنی، امکان بررسی‌های آماری گوناگون در مورد این کمیتها و برآورد فواصل اطمینان برای آنها را فراهم می‌آورد.

ماتریس کوواریانس مجهولات برآورده شده $(\hat{\underline{X}})$

با داشتن $\underline{C}_l$, $\hat{\underline{x}} = \underline{B}\,\underline{l}$ می‌خواهیم $\underline{C}_{\hat{x}} = \underline{B}\,\underline{C}_l\,\underline{B}^T$ بدست آوریم. با اعمال قانون انتشار خطاها به رابطه (۷) داریم:

$$C_{\hat{x}} = \left[(A^T P A)^{-1} A^T\right] C_l \left[(A^T P A)^{-1} A^T\right]^T \qquad \ldots(۲۹)$$

که در آن P برابر است با:

$$\underline{P} = \underline{Q}^{-1} = {}^2\underline{C}_l^{-1}$$

بنابر این (۲۹) را می‌توان به صورت ذیل نوشت:

$$\underline{C}_{\hat{x}} = (\underline{A}^T \underline{C}_l^{-1} \underline{A})^{-1} \underbrace{\underline{A}^T \underline{C}_l^{-1} \underline{A} (\underline{A}^T \underline{C}_l^{-1} \underline{A})^{-1}}_{\underline{I}}$$

$$= (\underline{A}^T \underline{C}_l^{-1} \underline{A})^{-1} \qquad \ldots (۲۹)$$

ماتریس کوواریانس تصحیحات برآورد شده $(\hat{\underline{v}})$

با توجه به مدل ریاضی $(\underline{v}+\underline{l}=\underline{A}\,\underline{x})$ داریم :

$$\hat{\underline{v}} = \underline{A}\,\hat{\underline{x}} - \underline{l}$$

$$= \underline{A}\,(\underline{A}^T\underline{P}\,\underline{A})^{-1}\,\underline{A}^T\underline{P}\,\underline{l} - \underline{l}$$

$$= \left[\underline{A}\,(\underline{A}^T\underline{P}\,\underline{A})^{-1}\underline{A}^T\underline{P} - \underline{I}\right]\underline{l}$$

بنابر این:

$$\underline{C}_{\hat{v}} = \left[\underline{A}\,(\underline{A}^T\underline{P}\,\underline{A})^{-1}\underline{A}^T\underline{P} - \underline{I}\right]\underline{C}_l\left[\underline{P}\,\underline{A}\,(\underline{A}^T\underline{P}\,\underline{A})^{-1}\underline{A}^T - \underline{I}\right]$$

با در نظر گرفتن:

$$\underline{P} = \underline{Q}^{-1} = \sigma_0^2\,\underline{C}_l^{-1}$$

داریم :

$$\underline{C}_{\hat{v}} = \left[\underline{A}\,(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T\underline{C}_l^{-1} - \underline{I}\right]\underline{C}_l\left[\underline{C}_l^{-1}\underline{A}(\underline{A}^T\underline{C}^{-1}\underline{A})^{-1}\underline{A}^T - \underline{I}\right]$$

$$= \left[\underline{A}\,(\underline{A}^T\underline{C}_l^{-1}\underline{A})\,\underline{A}^T - \underline{C}\right]\left[\underline{C}^{-1}\underline{A}(\underline{A}^T\underline{C}^{-1}\underline{A})^{-1}\underline{A} - \underline{I}\right]$$

$$= \underline{A}(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T\underline{C}_l^{-1}\underline{A}(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T - \underline{A}(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T \cancel{- \underline{A}(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}}$$

$$- \underline{A}\,(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T + \underline{C}_l = \underline{C}_l - \underline{A}\,(\underline{A}^T\underline{C}_l^{-1}\underline{A})^{-1}\underline{A}^T \qquad \ldots (۳۰)$$

ماتریس کوواریانس مشاهدات برآورده شده ($\hat{\underline{l}}$)

$$\hat{\underline{l}} = \underline{l} + \hat{\underline{v}} = \underline{A}\,\hat{\underline{x}}$$

$$\underline{C}_{\hat{l}} = \underline{A}\,\underline{C}_{\hat{x}}\,\underline{A}^T = \underline{A}\,(\underline{A}^T\,\underline{C}_l^{-1}\,\underline{A})^{-1}\,\underline{A}^T \qquad \ldots (۳۱)$$

با مقایسه (۳۰) و (۳۱) به این نتیجه می‌رسیم :

$$\underline{C}_{\hat{v}} = \underline{C}_l - \underline{C}_{\hat{l}} \qquad \ldots (۳۲)$$

نکته حائز اهمیت در ماتریسهای کوواریانس حاصله ، آن است که هیچکدام از آنها بستگی به مشاهدات و مقادیر آنها ندارد و همهٔ آنها را تنها با اختیار داشتن $\underline{C}_l$ و مدل ریاضی می‌توان بدست آورد .این مطلب از نقطه نظر طراحی بسیار با ارزش بوده و نکته مثبتی به شمار می‌رود . چون پیش از انجام مشاهدات می‌توان بررسی کرد که آیا دقت خواسته شده برای مجهولات قابل دستیابی است یا خیر . اما از طرفی دیگر در تمام این روابط $\underline{C}_l$ مورد نیاز است در حالی که در شروع سرشکنی غالبا $\underline{Q}$ در اختیار می‌باشد و رسیدن به برآوردی از $\underline{C}_l$ یکی از اهداف سرشکنی است.

خوشبختانه همانگونه که قبلا دیدید برآورد مقدار مجهولات نیازی به آگاهی از مقدار صحیح σ_0^2 ندارد .

بنابر این $\hat{\underline{x}}$، $\hat{\underline{l}}$ و $\hat{\underline{v}}$ بدون هیچگونه تاثیر پذیری از مقدار انتخابی برای σ_0^2 بدست می‌آیند . پس از برآورد این کمیتها اگر بتوانیم برآوردی برای σ_0^2 بدست آوریم گمان می‌رود که مشکل

ماتریسهای کوواریانس نیز حل گردد.

۲/۱۱- درجهٔ آزادی

یکی از نکات اساسی و مهم قبل از سرشکنی یافتن رابطهٔ بین مشاهدات و مجهولات است. به این رابطه همانطور که قبلا گفته شد مدل ریاضی می‌گوئیم. با در اختیار داشتن نحوهٔ ارتباط مشاهدات با مجهولات (مدل ریاضی) می‌توانیم مشخص کنیم که چند معادله برای تعیین مجهولات بصورت منحصر به فرد لازم است. اما همانگونه که پیشتر نیز گفته شد ما به این تعداد معادله اکتفا نمی‌کنیم چون به این ترتیب امکان پی‌بردن به اشتباه، برآورد مقدار قابل اطمینان برای مجهولات و برآورد دقت اندازه گیری میسر نخواهد بود. بنابراین تعداد معادلات را عمدا بیشتر از مجهولات قرار می‌دهیم. به این تعداد معادلات اضافی درجه آزادی گفته می‌شود (Degrees of freedom).

$$df = m - u \qquad \ldots (۳۳)$$

df : درجهٔ آزادی
m : تعداد معادلات
u : تعداد مجهولات

در مدل پارامتریک اگر تعداد معادلات را n بنامیم، درجه آزادی چنین خواهد بود:

$$df = n - u \qquad \ldots (۳۴)$$

علت بکار بردن لفظ آزادی (یا درجهٔ آزادی) بدین خاطر است که هر اندازه تعداد معادلات بیشتر از تعداد مجهولات گردد توان یا

آزادی عمل ما در پی بردن به اشتباهات و برآورد بهتر برای کمیتهای مجهول بالاتر خواهد رفت . در مدل پارامتریک ، تغییرات درجهء آزادی بین صفر تا تعداد مشاهدات است .

۲/۱۲ـ برآورد واریانس وزن واحد

همانگونه که قبلا گفته شد دستیابی به ماتریسهای کوواریانس $\hat{\underline{x}}$ ، $\hat{\underline{v}}$ و $\hat{\underline{l}}$ هنگامی میسر خواهد بود که $\underline{C}_l$ یا $\sigma_0^2 \underline{Q}$ در اختیار باشد . اما چنین چیزی عملا بندرت اتفاق می‌افتد . لذا می‌بایست بنحوی مقدار σ_0^2 را از طریق مشاهدات انجام شده برآورد کرد . یعنی به یک $\hat{\sigma}_0^2$ دست یافت و از آن به جای σ_0^2 موجود در روابط مربوط به ماتریسهای کوواریانس کمیتهای برآورد شده $(\hat{\underline{l}}, \hat{\underline{v}}, \hat{\underline{x}})$ استفاده کرد . اما این برآورد کننده (Estimator) بایستی ناریب باشد . نشان خواهیم داد که اولا

$$\hat{\sigma}_0^2 = \frac{\hat{\underline{v}}^T \underline{P}\, \underline{v}}{n-u}$$

و در ثانی این برآورد کننده ناریب است. یعنی :

$$E(\hat{\sigma}_0^2) = E\left[\frac{\hat{\underline{v}}^T \underline{P}\, \underline{v}}{n-u}\right] = \frac{1}{n-u} E(\underline{v}^T \underline{P}\, \underline{v}) = \sigma_0^2$$

یا

$$E(\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}}) = (n-u)\, \sigma_0^2 \qquad \ldots (۳۵)$$

اثبات

برای این منظور از معادلات نرمال شروع می کنیم :

$$\underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}} = \underline{A}^T \underline{P}\, \underline{l} \qquad (۳۶)$$

این دستگاه را می‌توان بصورت زیر نوشت:

$$\underline{A}^T \underline{P}\, (\underline{A}\, \hat{\underline{x}} - \underline{l}) = \underline{0} \quad \text{یا} \quad (\underline{A}\, \hat{\underline{x}} - \underline{l})^T \underline{P}\, \underline{A} = \underline{0} \qquad (۳۷)$$

$$\underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}} = \underline{A}^T \underline{P}\, \underline{l} \qquad \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A} = \underline{l}^T \underline{P}\, \underline{A}$$

بر اساس این تساوی می توان نشان داد :

$$\underline{v}^T \underline{P}\, \underline{v} - \hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}} = (\hat{\underline{x}} - \underline{x})\, \underline{A}^T \underline{P}\, \underline{A}\, (\hat{\underline{x}} - \underline{x})$$

چون :

$$\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}} = (\underline{A}\, \hat{\underline{x}} - \underline{l})^T \underline{P}\, (\underline{A}\, \hat{\underline{x}} - \underline{l})$$

$$= (\underline{A}\, \hat{\underline{x}} - \underline{l})^T \underline{P}\, \underline{A}\, \hat{\underline{x}} - \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{l} + \underline{l}^T \underline{P}\, \underline{l}$$

با توجه به (۳۷)

$$= \underline{l}^T \underline{P}\, \underline{l} - \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}} \qquad (۳۸)$$

و همچنین :

$$\underline{v}^T \underline{P}\, \underline{v} = (\underline{A}\, \underline{x} - \underline{l})^T \underline{P}\, (\underline{A}\, \underline{x} - \underline{l})$$

$$= \underline{x}^T \underline{A}^T \underline{P}\, \underline{A} - \underline{l}^T \underline{P}\, \underline{A}\, \underline{x} - \underline{x}^T \underline{A}^T \underline{P}\, \underline{l} + \underline{l}^T \underline{P}\, \underline{l}$$

اما $\underline{A}^T \underline{P}\, \underline{l} = \underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}}$ و $\underline{l}^T \underline{P}\, \underline{A} = \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A}$ بنابراین

$$= \underline{x}^T \underline{A}^T \underline{P}\, \underline{A}\, \underline{x} - \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A}\, \underline{x} - \underline{x}^T \underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}} + \underline{l}^T \underline{P}\, \underline{l} \qquad (۳۹)$$

بنابراین از تفاضل (۳۸) و (۳۹) خواهیم داشت :

$$\underline{v}^T \underline{P}\, \underline{v} - \hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}} = \underline{x}^T \underline{A}^T \underline{P}\, \underline{A}\, \underline{x} - \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A}\, \underline{x} - \underline{x}^T \underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}} + \underline{l}^T \underline{P}\, \underline{l} - \underline{l}^T \underline{P}\, \underline{l} + \hat{\underline{x}}^T \underline{A}^T \underline{P}\, \underline{A}\, \hat{\underline{x}}$$

$$= (\hat{\underline{x}} - \underline{x})^T \underline{A}^T \underline{P}\, \underline{A}\, (\hat{\underline{x}} - \underline{x})$$

در نتیجه :

$$\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}} = \underline{v}^T \underline{P}\, \underline{v} - (\hat{\underline{x}} - \underline{x})^T \underline{A}^T \underline{P}\, \underline{A}\, (\hat{\underline{x}} - \underline{x})$$

برای یک فرم مربعی(Quadratic) مانند $\underline{y}^T \underline{A}\, \underline{y}$ داریم:

$$\mathrm{Trace}(\underline{y}^T \underline{A}\, \underline{y}) = \underline{y}^T \underline{A}\, \underline{y}$$

و در مورد Trace آن:

$$\mathrm{Tr}(\underline{y}^T \underline{A}\, \underline{y}) = \mathrm{Tr}(\underline{y}(\underline{y}^T \underline{A}))$$

بنابراین:

$$\underline{y}^T \underline{A}\, \underline{y} = \mathrm{Tr}(\underline{y}\, \underline{y}^T \underline{A})$$

$$\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}} = \mathrm{Tr}(\underline{v}\, \underline{v}^T \underline{P}) - \mathrm{Tr}((\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T \underline{A}^T \underline{P}\, \underline{A}$$

$$= \mathrm{Tr}(\underline{v}^T \underline{v}\ \ {}^2\underline{C}_l^{-1}) - \mathrm{Tr}((\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T\ \ {}^2\underline{C}_{\hat{x}}^{-1})$$

حال به طرفین تساوی اخیر اپراتور E را اعمال می‌کنیم:

$$E(\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}}) = \sigma_0^2 E\,[\mathrm{Tr}(\underline{v}\, \underline{v}^T \underline{C}_l^{-1})] - \sigma_0^2 E\,[\mathrm{Tr}((\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T \underline{C}_{\hat{x}}^{-1})]$$

$$= \sigma_0^2 \mathrm{Tr}\,[E(\underline{v}\, \underline{v}^T \underline{C}_l^{-1})] - \sigma_0^2 \mathrm{Tr}\,[E((\hat{\underline{x}} - \underline{x})(\hat{\underline{x}} - \underline{x})^T \underline{C}_{\hat{x}}^{-1})]$$

$$= \sigma_0^2 (\mathrm{Tr}(\underline{C}_l\, \underline{C}_l^{-1}) - \mathrm{Tr}(\underline{C}_{\hat{x}}\, \underline{C}_{\hat{x}}^{-1}))$$

$$= \sigma_0^2 (\mathrm{Tr}\,(\underline{I}_n) - \mathrm{Tr}\,(\underline{I}_u))$$

$$= \sigma_0^2 (n - u)$$

بنابراین نشان دادیم که:

$$E(\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}}) = \sigma_0^2 (n - u)$$

و با توجه به رابطه (۳۵):

$$E(\hat{\sigma}_0^2) = \frac{E(\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}})}{n - u} = \sigma_0^2$$

یا به عبارتی دیگر:

$$\hat{\sigma}_0^2 = \frac{\hat{\underline{v}}^T \underline{P}\, \hat{\underline{v}}}{n - u} \qquad \ldots (۴۰)$$

یک برآورد کننده‌ٔ ناریب برای σ_0^2 است .

(n-u) همانگونه که قبلا گفته شد درجه‌ٔ آزادی سرشکنی است

(در مدل پارامتریک df = n-u) .

۲/۱۳- برآورد ماتریسهای کوواریانس کمیتهای حاصل از سرشکنی مدل پارامتریک خطی

پس از برآورد σ_0^2 ماتریسهای کوواریانس نیز از طریق برآورد کننده‌های ناریب ذیل قابل برآوردند .

$$\hat{\underline{C}}_{\hat{x}} = \hat{\sigma}_0^2 (\underline{A}^T \underline{P}\, \underline{A})^{-1} = \hat{\sigma}_0^2 (\underline{A}^T \underline{Q}^{-1} \underline{A})^{-1} \qquad \ldots (۴۱)$$

بطوری که :

$$E(\hat{\underline{C}}_{\hat{x}}) = \underline{C}_{\hat{x}}$$

$$\hat{\underline{C}}_{\hat{l}} = \underline{A}\, \hat{\underline{C}}_{\hat{x}}\, \underline{A}^T = \hat{\sigma}_0^2 \underline{A}(\underline{A}^T \underline{Q}^{-1} \underline{A})^{-1} \underline{A}^T \qquad \ldots (۴۲)$$

بطوریکه :

$$E(\hat{\underline{C}}_{\hat{l}}) = \underline{C}_{\hat{l}}$$

و بالاخره ،

$$\hat{\underline{C}}_{\hat{v}} = \hat{\sigma}_0^2 (\underline{Q} - \underline{A}(\underline{A}^T \underline{Q}^{-1} \underline{A})^{-1} \underline{A}^T) \qquad \ldots (۴۳)$$

و

$$E(\hat{\underline{C}}_{\hat{v}}) = \underline{C}_{\hat{v}}$$

<u>نکته</u> :

$\underline{C}_l$ ماتریس کوریانس مشاهدات است ، کمیتی که در سرشکنی معیار انتخاب وزن قرار می‌گیرد :

$$\underline{P} = \sigma_0^2 \underline{C}_l^{-1}$$

از طرفی دیگر بعد از سرشکنی می توان مشاهدات تصحیح شده (مشاهداتی که در مدل $\hat{\underline{l}} = \underline{A}\,\hat{\underline{x}}$ صدق می‌کنند) را بدست آورد .

$$\underline{C}_{\hat{l}} = \underline{A}\,\underline{C}_{\hat{x}}\underline{A}^T$$

دو تفاوت عمده بین $\underline{C}_l$ و $\underline{C}_{\hat{l}}$ وجود دارد ، اول آنکه $\underline{C}_{\hat{l}}$ (حتی اگر $\underline{C}_l$ قطری انتخاب شده باشد) دارای کوریانس (عناصر خارج از قطر اصلی) مخالف صفر است . دوم آنکه $\underline{C}_{\hat{l}}$ مجرد است ($|\underline{C}_{\hat{l}}| = \underline{0}$) چون ماتریسی nxn (n = تعداد مشاهدات) و از مرتبهٔ u (rank) (تعداد مجهولات) است . چون مرتبهٔ $\underline{A}$ و $\underline{C}_{\hat{x}}$ برابر u است لذا $\underline{C}_{\hat{l}}$ نمی‌تواند معیار انتخاب وزن قرار گیرد .

۲/۱۴- مدل پارامتریک غیر خطی

فرم کلی مدل ریاضی به صورت $\underline{f}(\underline{x},\underline{l})=\underline{0}$ است . حالت خاصی از آن ، یعنی $\underline{l}=\underline{f}(\underline{x})$ ، را مدل پارامتریک (غیر خطی) می‌نامیم (پارامتریک بدین خاطر آن که از حل این مدل مستقیما به پارامترهای مجهول می‌رسیم و مشاهده به لحاظ صراحت آن نسبت به مشاهدات) . این مدل در واقع دستگاهی است با تعداد معادلات (مشاهدات) بیش از مجهولات می‌باشد ، که به خاطر تصادفی بودن مشاهدات ، ناسازگار و بدون جواب است . اگر به این مدل مجهولات جدیدی به صورت زیر اضافه کنیم ،

$$\underline{v} + \underline{l} = \underline{f}(\underline{x}) \qquad \ldots (۲)$$

این دستگاه به دستگاهی با بینهایت جواب تبدیل می‌گردد . حال از بین این بینهایت جواب ، جوابی را بر می‌گزینیم که برای آن شرط زیر برقرار باشد .

$$w = \underline{v}^T \underline{P}\, \underline{v} = \min. \qquad \ldots (۳)$$

برای رسیدن به این جواب از w نسبت به x مشتق گرفته مساوی صفر قرار می‌دهیم .

$$w = \underline{v}^T \underline{P}\, \underline{v}$$

$$= (\underline{f}(\underline{x}) - \underline{l})^T \underline{P} (\underline{f}(\underline{x}) - \underline{l}) \qquad \text{(از تساوی (۲))}$$

$$\frac{dw}{dx} = \frac{dw}{d(\underline{f}(\underline{x}) - \underline{l})} \cdot \frac{d(\underline{f}(\underline{x}) - \underline{l})}{d\underline{x}}$$

$$= 2(\underline{f}(\underline{x}) - \underline{l})^T \underline{P} \frac{d(\underline{f}(\underline{x}))}{d\underline{x}} = \underline{0}$$

که در آن :

$$\frac{d\,(\,\underline{f}(\underline{x})\,)}{d\underline{x}} = \begin{bmatrix} \frac{d f_1(x)}{dx_1} & \frac{d f_1(x)}{dx_2} & \cdots & \frac{d f_1(x)}{dx_u} \\ \frac{d f_2(x)}{dx_1} & \frac{d f_2(x)}{dx_2} & \cdots & \frac{d f_2(x)}{dx_u} \\ \vdots & \vdots & & \vdots \\ \frac{d f_n(x)}{dx_1} & \frac{d f_n(x)}{dx_2} & & \frac{d f_n(x)}{dx_u} \end{bmatrix} = \underline{A}$$

بنابر این خواهیم داشت :

$$2\,(\,\underline{f}(\underline{x}) - \underline{l}\,)^T\,\underline{P}\,\underline{A} = \underline{0}$$

یا

$$\underline{f}(\underline{x})^T\,\underline{P}\,\underline{A} - \underline{l}^T\,\underline{P}\,\underline{A} = \underline{0} \qquad \ldots (۴)$$

دستگاه فوق یک دستگاه u معادله‌ای u مجهولی غیر خطی است . به این دستگاه به خاطر برابری تعداد معادلات با مجهولات دستگاه معادلات نرمال می‌گوییم . در ریاضیات هیچ روش تحلیلی کلی برای حل دستگاه‌های معادلات خطی وجود ندارد . اما حل عددی آنها با بسط معادلات حول مقادیر تقریبی مجهولات به سری تیلور ، صرف نظر کردن از جملات غیر خطی بسط ، ایجاد دستگاه معادلات خطی و تکرار بسط حول مقدار جدید بدست آمده برای مجهولات مسیر خواهد بود .

بنابر این چون می‌دانیم مدل پارامتریک غیر خطی نهایتا به دستگاه معادلات نرمال غیر خطی خواهد رسید ، مدلهای غیر خطی را در ابتدا با بسط به سری تیلور و صرف نظر کردن از جملات بالاتر از مرتبهٔ یک خطی می‌کنیم . به این ترتیب به معادلات نرمال خطی خواهیم رسید که از طریق روشهای معمول حل دستگاه‌های معادلات معادلات خطی قابل حل است.

این عمل باعث می‌گردد که بتوان یک روش حل کلی برای مدلهای پارامتریک خطی و غیر خطی بدست آورد .

اگر (۲) را حول $\underline{x}^0$ به سری تیلور بسط دهیم و جملات غیر خطی را کنار بگذاریم ، خواهیم داشت :

$$\underline{v} + \underline{l} = \underline{f}(\underline{x}^0) + \frac{d\,\underline{f}}{d\,\underline{x}}(\underline{x} - \underline{x}^0) \qquad \ldots (۵)$$

و با فرض

$$\left[\begin{array}{l} \underline{x} - \underline{x}^0 = \underline{x} \\ \underline{l} - \underline{f}(\underline{x}^0) = \underline{l} \\ \dfrac{d\,\underline{f}}{d\,\underline{x}} = \underline{A} \end{array}\right.$$

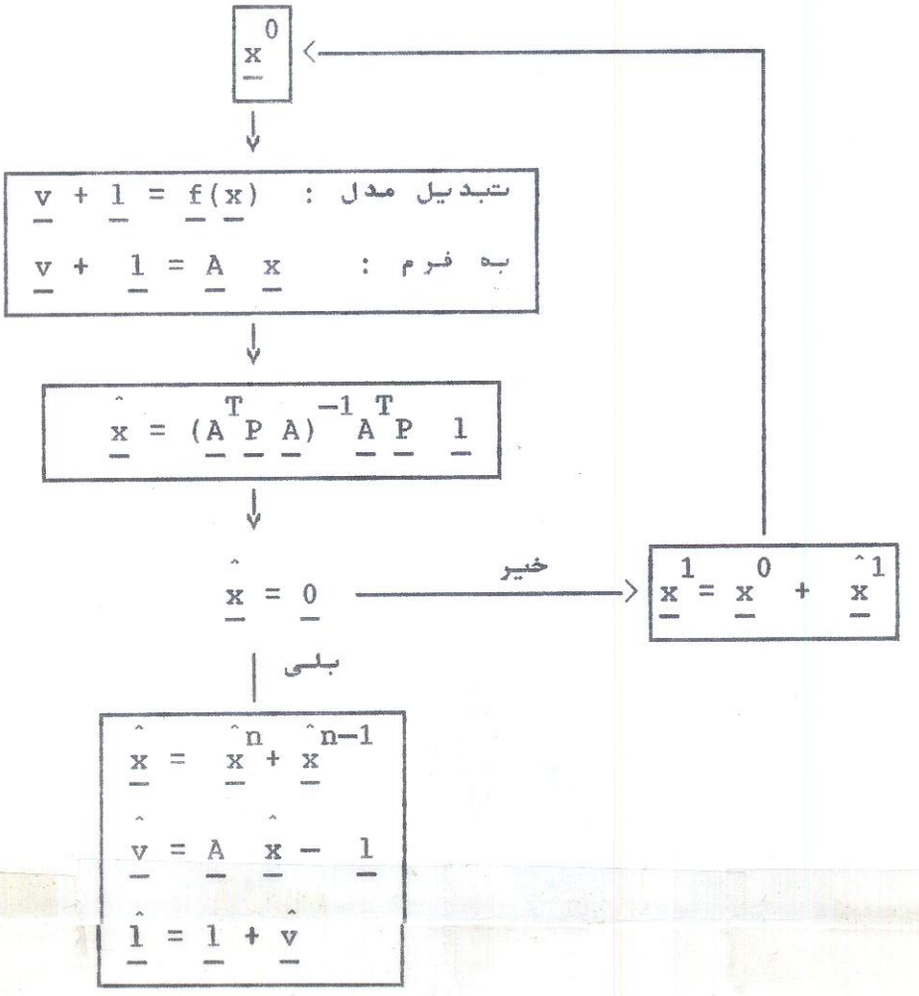
به رابطهٔ ذیل می‌رسیم ،

$$\underline{v} + \underline{l} = \underline{A}\ \underline{x} \qquad \ldots (۶)$$

که دقیقا مشابه رابطهٔ (۲) است. بنابر دارای جواب کمترین مربعاتی مشابه (۷) است.

۲/۱۵- نحوهٔ تکرار و تعیین پارامترهای مجهول در مدل پارامتریک غیر خطی

در بخش قبل ملاحظه فرمودید که چگونه می‌توان مدل پارامتریک غیر خطی را به فرمی مشابه معادلات پارامتریک خطی در آورد. که نهاینا در راه رسیدن به جواب کمترین برای آن به دستگاههای معادلات نرمال خطی خواهیم رسید که از طریق آنها مقادیر مجهولات تعیین خواهد شد. چون مقادیری که بسط حول آن صورت گرفته تقریبی بوده می‌توان مقدار جدید مجهولات را به عنوان نقطهٔ بسط جدید بکار برده و مجددا مدل را خطی و به جواب کمترین مربعات آن رسید. این عمل می‌تواند آنقدر تکرار گردد تا جواب حاصل برای جزئیات این تکرار در شکل ذیل نشان داده شده است.



در تکرار مرحلهء آخر $\hat{\underline{l}}$ را می‌توان از رابطهء خطی $\hat{\underline{l}}=\underline{A}\hat{\underline{x}}$ نیز بدست آورد ، چون:

$$\hat{\underline{l}} = \underline{l} + \hat{\underline{v}} = \underline{l} + \underbrace{\underline{A}\ \hat{\underline{x}} - \underline{l}}_{\hat{\underline{v}}} = \underline{l} + \underline{A}\ \hat{\underline{x}} - \underbrace{(\underline{l} - \underline{f}(\underline{x}^{n-1}))}_{\underline{l}}$$

$$= \underline{A}\ \hat{\underline{x}} + \underline{f}(\underline{x}^{n-1}) = \underline{A}\ \hat{\underline{x}} + \underline{A}\ \underline{x}^{n-1} = \underline{A}(\hat{\underline{x}} + \underline{x}^{n-1}) = \underline{A}\ \hat{\underline{x}} \qquad \ldots (۴۴)$$

تمرین :

ماتریس کوواریانس کمیتهای حاصل از سرشکنی مدل پارامتریک غیر خطی را بدست آورید .

۲/۱۶- سرشکنی معادلات شرط

قبلا گفتیم که اگر تعداد معادلات بیشتر از مجهولات باشد ، درجه آزادی خواهیم داشت . در صورت وجود درجهء آزادی ، به تعداد آن میتوان معادله به فرم:

$$\underline{g}(\underline{l}) = \underline{0} \qquad \ldots (۴۵)$$

بین مشاهدات ترتیب داد . به این معادلات که تعداد آنها برابر درجهء آزادی است معادلات شرط گفته می‌شود (Condition Equations) . این معادلات معمولا روابط هندسی موجود بین مشاهدات هستند ، که به خاطر خطای مشاهدات (تصادفی بودن مشاهدات) هیچگاه

برقرار نخواهد بود (فرم خطی این معادلات به صورت $\underline{B}\,\underline{l} = \underline{c}$ قابل نمایش است) .

با توجه به خطای اتفاقی مشاهدات به خود اجازه می‌دهیم که به مشاهدات تصحیحاتی کرده تا معادلات سازگار شوند . به این ترتیب (در فرم خطی) خواهیم داشت:

$$\underline{B}\,(\,\underline{l} + \underline{v}\,) = \underline{C} \qquad \ldots (۴۶)$$

این تصحیحات مجهول‌اند و چون همیشه تعداد مشاهدات کمتر از درجهٔ آزادی است دستگاه معادلات (۴۶) بینهایت جواب خواهد داشت از بین بی‌نهایت جواب موجود جوابی را انتخاب کنیم که علاوه بر معادلات (۴۶) در شرط

$$\underline{v}^T\,\underline{P}\,\underline{v} = \min \qquad \ldots (۴۷)$$

صدق کند که جوابی منحصر به فردی برای $\underline{v}$ است (در صورت کمبود مرتبه نداشتن $\underline{B}$) و به آن جواب کمترین مربعات یا برآورد کمترین مربعات اطلاق می‌گردد .

از معادلهٔ (۴۶) داریم :

$$\underline{B}\,\underline{l} + \underline{B}\,\underline{v} = \underline{C} \qquad \text{یا} \qquad \underline{B}\,\underline{l} - \underline{C} + \underline{B}\,\underline{v} = \underline{0}$$

$\underline{w} = \underline{B}\,\underline{l} - \underline{C}$ را اختلاف بست (Misclosuer) می‌نامیم چون اگر معادلات خطایی نداشتند $\underline{w}$ مساوی صفر می‌بود . بنابراین تفاوت آن از صفر را می‌توان خطای در بسته شدن تساوی دانست . بنابراین معادلات شرط بعد از اعمال تصحیحات مجعول بفرم زیر در خواهند آمد :

$$\underline{B}\,\underline{v} + \underline{w} = \underline{0} \qquad \ldots (۴۸)$$

حال معادلات شرط غیر خطی را در نظر می‌گیریم . با بسط به سری تیلور این معادلات حول $\underline{l}$ (مقادیر مشاهدات) داریم :

$$\underbrace{g(\underline{l})}_{\underline{w}} + \underbrace{\frac{g}{\underline{l}}}_{\underline{B}} \underbrace{(\hat{\underline{l}} - \underline{l})}_{\underline{v}} = \underline{0}$$

که آن را می توان به فرم زیر نوشت:

$$\underline{B}\,\underline{v} - \underline{w} = \underline{0}$$

ملاحظه می‌گردد که در اینگونه مدلها نیز پس از خطی کردن به فرم $\underline{B}\,\underline{v} + \underline{w} = \underline{0}$ خواهیم رسید . پس بقیهٔ مراحل برای مدلهای خطی و غیر خطی یکسان خواهد بود . برای بدست آوردن جوابی که هم در شرط (۴۷) صدق کند و هم در معادلهٔ (۴۸) از روش لاگرانژ استفاده می‌کنیم .

$$\phi = \underline{v}^T \underline{P}\,\underline{v} + 2\,\underline{k}^T \underbrace{(\underline{B}\,\underline{v} + \underline{w})}_{=0}$$

$$\left[\begin{array}{ll} \dfrac{\partial \phi}{\partial \underline{v}} = 2\,\underline{v}^T \underline{P} + 2\,\underline{k}^T \underline{B} = \underline{0} & \ldots (۴۹) \\[2ex] \dfrac{\partial \phi}{\partial \underline{k}} = 2\,(\underline{B}\,\underline{v} + \underline{w})^T = \underline{0} & \ldots (۵۰) \end{array}\right.$$

دستگاه فوق دستگاه معادلات نرمال است که از حل آن به $\hat{\underline{k}}$ و $\hat{\underline{v}}$ دست خواهیم یافت .

با توجه به (۴۹) :

$$\underline{P}\,\underline{v} = -\underline{B}^T \underline{k}$$

$$\Rightarrow \quad \hat{\underline{v}} = -\underline{P}^{-1}\underline{B}^{T}\underline{k} \qquad \ldots (۵۱)$$

و با توجه به (۵۰) و (۵۱)

$$-\underline{B}\,\underline{P}^{-1}\underline{B}^{T}\hat{\underline{k}} + \underline{w} = 0$$

یا

$$(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})\,\hat{\underline{k}} = \underline{w}$$

در صورت غیر مجرد بودن $(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})$

$$\hat{\underline{k}} = (\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{w} \qquad \ldots (۵۲)$$

و بالاخره با استفاده ء از (۵۱) و (۵۲)

$$\hat{\underline{v}} = -\underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{w} \qquad \ldots (۵۳)$$

<u>توجه</u>: $\hat{\underline{v}}$ و $\hat{\underline{k}}$ مقادیر خاصی از $\underline{v}$ و $\underline{k}$ هستند که در شرط (۴۷) و معادلات (۴۸) صدق می‌کنند .

<u>نکته</u>: در اینجا نیز اگر بجای $\underline{P}^{-1}$ از $\underline{Q}$ یا $\underline{C}_{-1}$ استفاده شود نتیجه برآورد یکسان خواهد بود .

برای اینکه نشان دهیم این جواب مینیمم کننده است از رابطه ء (۴۹) نسبت به $\underline{v}$ مجددا مشتق می‌گیریم :

$$\frac{\partial^2 \phi}{\partial \underline{v}^2} = 2\,\underline{P} \qquad \ldots (۵۴)$$

بنابراین در صورتیکه $\underline{P}$ را اکیدا مثبت انتخاب کرده باشیم شرط

(۴۷) به ازاء مقادیر برآورده شده برقرار خواهد بود .

۲/۱۷- جواب حاصل از سرشکنی معادلات شرط و شبه معکوس

در بخش قبل دیدید که مدل :

$$\underline{B}\,\underline{v} + \underline{w} = \underline{0}$$

دستگاهی با تعداد مجهولات ($\underline{v}$) بیشتر از معادلات است و با شرط کمترین مربعات یعنی:

$$\underline{v}^T \underline{P}\,\underline{v} = \min$$

به جواب منحصر به فرد زیر رسیدیم :

$$\hat{\underline{v}} = -\underline{P}^{-1}\underline{B}^T(\underline{B}\,\underline{P}^{-1}\underline{B}^T)^{-1}\underline{w}$$

یا

$$\hat{\underline{v}} = \underline{H}\,\underline{W}$$

یعنی $\underline{H}$ را می‌توانیم نوعی معکوس برای $\underline{B}$ دانست . از طرفی دیگر می‌دانیم که برای هر ماتریس دلخواه مانند $\underline{B}$ ماتریسی مانند $\underline{H}$ با خواص ذیل وجود دارد .

(i) $\underline{B}\,\underline{H}\,\underline{B} = \underline{B}$

(اگر $\underline{H}$ تنها دارای این خاصیت باشد معکوس فراگیر (Generalized Inverse) نامیده می‌شود . برای هر ماتریس بینهایت معکوس فراگیر می‌توان یافت. اگر $\underline{H}$ علاوه بر شرط (i) در شروط ذیل نیز صدق کند شبه معکوس (Pseudo Inverse) نامیده می‌شود .

(ii) $\underline{H}\,\underline{B}\,\underline{H} = \underline{H}\,\underline{B}$

(iii) $(\underline{H}\ \underline{B})^T = \underline{H}\ \underline{B}$

(iv) $(\underline{B}\ \underline{H})^T = \underline{B}\ \underline{H}$

شبه معکوی بر خلاف معکوس فراگیر منحصر به فرد است. حال این فکر به ذهن متبادل می‌گردد که چه ارتباطی بین $\underline{H}$ حاصل از روش کمترین مربعات و شبه معکوس $\underline{B}$ وجود دارد . در نظر داشته باشید که هر دو منحصر به فرد هستند. بنابر این بد نیست که چهار خاصیت فوق را برای ماتریس $\underline{H}$ حاصل از روش کمترین مربعات بررسی کنیم. شرط (i) ، (ii) و (iv) به وضوح صادق‌اند. اما در مورد شرط سوم داریم:

$$\text{(iii)}\quad (\underline{H}\ \underline{B})^T = (\underline{P}^{-1}\underline{B}^T(\underline{B}\ \underline{P}^{-1}\underline{B}^T)^{-1}\underline{B})^T$$

$$= \underline{B}^T(\underline{B}\ \underline{P}^{-1}\underline{B}^T)^{-1}\underline{B}\ \underline{P}^{-1}$$

این شرط تنها در صورتی برقرار خواهد بود که $\underline{P}=\underline{I}$ باشد.

بنابراین در صورتیکه $\underline{P} = \underline{I}$ باشد $\underline{H}$ شبه معکوس $\underline{B}$ خواهد بود $\underline{H} = \underline{B}^+$.

از طرف دیگر $\underline{H}$ دارای خصوصیت جالب دیگری است:

$$\underline{B}\ \underline{H} = \underline{I}$$

یعنی اگر $\underline{H}$ را از سمت راست در $\underline{B}$ ضرب کنیم به ماتریس یکه خواهیم رسید، به این خاطر به $\underline{H}$ (Left Orthogonal Inverse) گفته می‌شود .(علت بکار بردن لفظ Orthogonal بعدا روشن می‌شود).

۲/۱۸ـ ماتریسهای کوواریانس کمیتهای برآورد شده از طریق سرشکنی مدل شرط

۱ـ ماتریس کوواریانس اختلاف بست ($\underline{w}$)

بردار اختلاف بست در حالت کلی بدین صورت بدست می‌آید :

$$\underline{w} = \underline{f}(\underline{l}_0)$$

در این رابطه $\underline{l}_0$ بردار مشاهدات و متغیر تصادفی است، بنابر این طبق قانون انتشار خطاها داریم :

$$\underline{C}_w = \frac{\partial \underline{f}}{\partial \underline{l}} \underline{C}_{l_0} \left(\frac{\partial \underline{f}}{\partial \underline{l}}\right)^T$$

$$= \underline{B}\, \underline{C}_l\, \underline{B}^T \qquad \ldots (۵۵)$$

۲ـ ماتریس کوواریانس تصحیحات برآورده شده $\hat{\underline{v}}$

$$\hat{\underline{v}} = -\underline{P}^{-1} \underline{B}^T (\underline{B}\, \underline{P}^{-1} \underline{B}^T)^{-1} \underline{w}$$

$$\underline{C}_{\hat{v}} = \underline{P}^{-1} \underline{B}^T \underbrace{(\underline{B}\, \underline{P}^{-1} \underline{B}^T)^{-1} \underline{C}_w (\underline{B}\, \underline{P}^{-1} \underline{B}^T)^{-1}}_{\underline{I}} \underline{B}\, \underline{P}^{-1}$$

$$= {}_0\underline{P}^{2\,-1} \underline{B}^T (\underline{B}\, \underline{P}^{-1} \underline{B}^T)^{-1} \underline{B}\, \underline{P}^{-1}$$

$$= \underline{C}_l\, \underline{B}^T (\underline{B}\, \underline{C}_l\, \underline{B}^T)^{-1} \underline{B}\, \underline{C} \qquad \left(P = \sigma_0^2 C_l^{-1} \Longrightarrow P^{-1} = \frac{C_l}{\sigma_0^2}\right) \qquad \ldots (۵۶)$$

ماتریس کوواریانس مشاهدات برآورد شده $(\hat{\underline{l}})$

$$\hat{\underline{l}} = \underline{l} + \hat{\underline{v}}$$

$$= \underline{l} - \underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{W}$$

$$= \underline{l} - \underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{f}(\underline{l})$$

$$\underline{C}_{\hat{l}} = (\underline{I} - \underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{B})\,\underline{C}_{l}(\underline{I} - \underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{B}\,\underline{P}^{-1})$$

$$= \underline{C}_{l} - \underline{C}_{l}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{B}\,\underline{P}^{-1} - \underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{B}\,\underline{C}_{l}$$

$$+ \underline{P}^{-1}\underline{B}^{T}\underbrace{(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{B}\,\underline{C}_{l}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}}_{\underline{I}}\underline{B}\,\underline{P}^{-1}$$

$$= \underline{C}_{l} - \underline{C}_{l}\underline{B}^{T}(\underline{B}\,\underline{C}_{l}\underline{B}^{T})^{-1}\underline{B}\,\underline{C}_{l}^{-1} = \underline{C}_{l} - \underline{C}_{\hat{v}} \qquad \ldots (۵۷)$$

نکته : با توجه به جواب کمترین مربعات حاصل برای $\hat{\underline{v}}$:

$$\hat{\underline{v}} = -\underline{P}^{-1}\underline{B}^{T}(\underline{B}\,\underline{P}^{-1}\underline{B}^{T})^{-1}\underline{w}$$

ملاحظه می گردد که حتی اگر ماتریس وزن را مساوی معکوس ماتریس کوواریانس نسبی قرار دهیم نتیجهء حاصل همان خواهد بود که در صورت انتخاب وزن بصورت عکس ماتریس کوواریانس حاصل می‌شود . اما در مورد ماتریسهای کوواریانسهای حاصل وضع چنین نیست و تمامی آنها در صورتی قابل محاسبه خواهند بود که ماتریس کوواریانس مشاهدات در دست باشد که معمولا چنین نیست . بدین خاطر سراغ برآورد کننده‌های ناریب آنها یعنی:

$$\hat{\underline{C}}_{\hat{v}} = \hat{\sigma}_0^2 \underline{Q}\, \underline{B}^T (\underline{B}\, \underline{Q}\, \underline{B}^T)^{-1} \qquad \ldots (۵۸)$$

$$\hat{\underline{C}}_{\hat{l}} = \hat{\sigma}_0^2 (\underline{Q} - \underline{Q}\, \underline{B}^T (\underline{B}\, \underline{Q}\, \underline{B}^T)^{-1} \underline{B}\, \underline{Q}) \qquad \ldots (۵۹)$$

$$\hat{\underline{C}}_{\hat{W}} = \hat{\sigma}_0^2 \underline{B}\, \underline{Q}\, \underline{B}^T \qquad \ldots (۶۰)$$

خواهیم رفت .

تمرین : ثابت کنید کمیتهای فوق‌الذکر برآورد کننده‌های نااریبی برای $\hat{\underline{C}}_{\hat{v}}$ ، $\hat{\underline{C}}_{\hat{l}}$ و $\hat{\underline{C}}_{\hat{W}}$ هستند .

نکته :

معادلهٔ شرط غیر خطی به فرم $\underline{f}(\underline{l}) = \underline{0}$ است که با بسط آن به سری تیلور حول مشاهدات و صرف نظر کردن از جملات غیر خطی به فرم مقابل تبدیل می‌کنیم : $\underline{B}\, \underline{v} + \underline{w} = \underline{0}$

نکتهٔ مهم در این بسط آن است که $\hat{\underline{l}}$ و $\underline{l}$ بسیار بهم نزدیک خواهند بود در نتیجه تقریب خطی تقریبی بسیار خوب از مدل در همسایگی $\underline{l}$ است ولذا نیازی به تکرار نیست و از مدل خطی شده می‌توان به جواب رسید .

۲/۱۹ـ تبدیل مشاهدات با وزن P به مشاهدات با وزن یکه (I)

در بخش (۱/۱۸) ذکر شد که سریعترین و دقیق‌ترین روش یافتن معکوس ماتریسهای اکیدا مثبت و متقارن روش تجزیهٔ چولسکی است. از طرفی دیگر در بخشهای قبلی این فصل ملاحظه کردید که ماتریس ضرایب معادلات نرمال ، ماتریسی اکیدا مثبت و متقارن است . بنابراین بهترین روش برای محاسبهٔ معکوس آن روش Choleski است .

بعلاوه از این روش می‌توان در تبدیل مشاهدات با وزن $\underline{P}$ به مشاهدات باوزن یکه استفاده کرد .

ثابت خواهیم کرد که اگر $\underline{l}$ دارای وزن $\underline{P}$ باشد $\underline{l}^* = \underline{L}^T \underline{l}$ ($\underline{P} = \underline{L}\,\underline{L}^T$ و $\underline{L}$ ماتریس پایین مثلثی است) دارای وزن یکه ($\underline{I}$) است یعنی :

$$\hat{\underline{x}} = (\underline{B}^T \underline{B})^{-1} \underline{B}^T \underline{l}^*$$

که B و l^* از معادلهٔ مشاهدهٔ اصلی ($\underline{l} = \underline{A}\,\underline{x}$) با ضرب طرفین از چپ در $\underline{L}^T$ حاصل شده اند :

$$\underbrace{\underline{L}^T \underline{l}}_{\underline{l}^*} = \underbrace{\underline{L}^T \underline{A}\,\underline{x}}_{\underline{B}\,\underline{x}}$$

$$\Longrightarrow \hat{\underline{x}} = (\underline{B}^T \underline{B})^{-1} \underline{B}^T \underline{l}^* = (\underline{A}^T \underbrace{\underline{L}\,\underline{L}^T}_{\underline{P}} \underline{A})^{-1} \underline{A}^T \underbrace{\underline{L}\,\underline{L}^T}_{\underline{P}} \underline{l} = (\underline{A}^T \underline{P}\,\underline{A})^{-1} \underline{A}^T \underline{P}\,\underline{l}$$

منابع

1, Cooper, M. A. R. Fundamentals of Survey Measurements and analysis (1974), Collins Technical Books.

2, Cooper, M. A. R. Control Surveys In Civil Engineering.

3, Vanicek, P. and Krakiwisky, E. Geodesy (1986), North Holland.

4, Mikhail, E. M. Observations and Least Squares (1976), IEP– A Dun-Donnelley Publisher, New York.

5, Krakiwisky, E. Papers For the CISM Adjustment and Analysis Seminars, (1987), The Canadian Institute of Surveying and Mapping, Ottawa, Canada.