FACULDADE · DE · CIÊNCIAS UNIVERSIDADE · DE · LISBOA

Observações em Ciências Geográficas: Métodos de Ajustamento e Análise

V. B. Mendes

1. INTRODUÇÃO

A compreensão de um sistema tão complexo como a Terra, bem como a sua interacção com outros planetas, requer a recolha, processamento e análise de dados terrestres e espaciais [AGU, 2004]. Embora a utilização destes dados esteja tradicionalmente associada à Topografia e Geodesia, um grande número de disciplinas (Fotogrametria, Detecção Remota, posicionamento com sistemas de posicionamento global, modelação digital do terreno, Sistemas de Informação Geográfica, processamento e visualização digital de imagem, cartografia computacional) utiliza, hoje em dia, dados geoespaciais. A combinação destas disciplinas constitui a Geomática ou Geoinformática, que se define como a Ciência e a Tecnologia para obter, analisar, interpretar, utilizar, arquivar e distribuir informação geoespacial.

Os problemas associados à recolha de dados, à sua manipulação matemática e análise não são exclusivos das Ciências da Terra e as metodologias que serão descritas podem ser aplicadas a outros domínios científicos.

A metodologia do processo que conduz à obtenção de dados, processamento de dados, análise e representação gráfica de resultados envolve, de um modo geral, diversas fases [Vanícek and Krakiwsky, 1986; Gracie and Krakiwsky, 1987]:

- *identificação dos parâmetros desconhecidos (incógnitas)*

Tal como noutros ramos da Ciência, a Engenharia Geográfica utiliza quantidades que são conhecidas e isentas de erro, **constantes**, e quantidades desconhecidas, que pretendemos determinar e que, na generalidade dos casos, não podem ser medidas directamente, **parâmetros desconhecidos**. As **observáveis** (**mensurandas**) são quantidades cuja classificação se situa entre aquelas entidades. Na terminologia aqui utilizada, uma **observável** é uma quantidade física ou geométrica que pode ser observada, ou seja, à qual se pode atribuir um valor numérico, **observação** (**valor**

medido), através de um **procedimento de medição**, utilizando um dado **instrumento** ou **sensor**.

Como exemplos de "constantes" (**valores convencionalmente verdadeiros**) podemos citar a velocidade de propagação da luz no vácuo e a constante gravitacional de Newton. Como exemplos de parâmetros desconhecidos usualmente pretendidos, podemos citar coordenadas de pontos, desvios da vertical e variações temporais destes parâmetros. Finalmente, como exemplos de observáveis podemos referir ângulos horizontais e direcções (observados com um teodolito), distâncias espaciais entre dois pontos (medidas quer por distanciómetros, quer por diversas técnicas espaciais), distâncias espaciais entre um ponto à superfície da Terra e um ponto num satélite (a partir de medições laser), diferenças de gravidade e gravidades absolutas (obtidas por gravímetros relativos e absolutos, respectivamente).

- *formulação do modelo matemático*

O **modelo matemático** estabelece a relação entre as observações e os parâmetros pretendidos e pode ser do tipo explícito ou implícito. O modelo explícito apresenta duas variantes: modelo directo e modelo indirecto. Qualquer modelo pode ainda ser do tipo linear ou não linear.

No **modelo directo**, os parâmetros desconhecidos são expressos em função das observações. Designando por $\boldsymbol{x}$ o vector dos parâmetros desconhecidos e por $\boldsymbol{\ell}$ o vector das observações, a representação matemática do modelo directo é do tipo

$$\boldsymbol{x} = \mathbf{b}(\boldsymbol{\ell}), \tag{1.1}$$

onde $\mathbf{b}$ é um vector de funções.

A Equação (1.1) pode ser escrita na versão linearizada

$$\boldsymbol{x} = \mathbf{B}\boldsymbol{\ell} + \boldsymbol{w}, \tag{1.2}$$

onde $\mathbf{B}$ é a matriz dos coeficientes das observações, que relaciona as observações com os parâmetros, e $\boldsymbol{w}$ é o vector de constantes, que reflecte quantidades físicas ou geométricas conhecidas. Como exemplos básicos de modelos directos pode-se referir:

(1) a medição de um comprimento e (2) o cálculo da área de um rectângulo, dadas as observações do comprimento dos seus lados. Para o primeiro exemplo, o modelo linear (1.1) fica reduzido à forma:

$$\boldsymbol{x} = \boldsymbol{\ell} \tag{1.3}$$

Sob certas condições, o modelo poderá não possuir quaisquer parâmetros desconhecidos. Este caso particular constitui o **modelo condicional**, uma vez que reflecte "condições" geométricas que relacionam entre si apenas observações. O modelo explícito tem neste caso a forma

$$\mathbf{b}(\boldsymbol{\ell}) = \mathbf{0} \,. \tag{1.4}$$

A versão linear correspondente será

$$\mathbf{B}\boldsymbol{\ell} + \boldsymbol{w} = \mathbf{0} \,. \tag{1.5}$$

As redes de nivelamento constituem exemplos típicos de modelos condicionais.

No **modelo indirecto** as observações são expressas em função dos parâmetros desconhecidos:

$$\boldsymbol{\ell} = \mathbf{a}(\boldsymbol{x}) \;, \tag{1.6}$$

onde **a** representa um vector de funções.

Na sua versão linear, tem-se

$$\boldsymbol{\ell} = \mathbf{A}\boldsymbol{x} + \boldsymbol{w} \;, \tag{1.7}$$

onde **A** é a matriz que relaciona os parâmetros desconhecidos com as observações. Contrariamente aos modelos directos, os modelos indirectos só em situações muito especiais conduzem a uma solução única. De um modo geral, os modelos são indeterminados e é necessário efectuar um **ajustamento**.

Um exemplo típico de um modelo indirecto é o cálculo das coordenadas de 2 pontos, a partir de observações de distâncias entre eles (no espaço, por exemplo):

$$d_{ij} = \sqrt{\left(x_j - x_i\right)^2 + \left(y_j - y_i\right)^2 + \left(z_j - z_i\right)^2} \,.$$

Obviamente uma observação única de distância não permite a determinação dos 6 parâmetros, pelo que será necessário estabelecer um conjunto de equações. Salienta-se ainda que este é também um caso típico de um modelo do tipo não linear.

No **modelo implícito** existe uma interligação mútua entre as observáveis e os parâmetros desconhecidos, pelo que as observáveis não podem ser expressas, de forma explícita, em função dos parâmetros, como acontecia nos modelos anteriores. Neste caso, o modelo é do tipo

$$\mathbf{f}\left(\boldsymbol{x},\boldsymbol{\ell}\right)=\mathbf{0} \ . \tag{1.8}$$

Em linguagem matricial, o modelo implícito escreve-se sob a forma:

$$\mathbf{A}\boldsymbol{x}+\mathbf{B}\boldsymbol{\ell}+\boldsymbol{w}=\mathbf{0} \ , \tag{1.9}$$

onde **A** é a **primeira matriz de configuração** e **B** é a **segunda matriz de configuração**. Facilmente se reconhece que os modelos explícitos (directo e indirecto) não são mais do que casos particulares de um modelo implícito.

Um modelo matemático compreende duas componentes [Mikhail, 1976]: o **modelo funcional**, que descreve as propriedades determinísticas da situação física ou geométrica considerada, e o **modelo estocástico**, que designa e descreve as propriedades estocásticas (probabilísticas) das variáveis envolvidas no modelo funcional. Tomando o exemplo de um triângulo no plano, suponha-se que os seus ângulos internos foram medidos e que se têm observações redundantes (ou seja, em número superior ao estritamente necessário). A solução do problema depende do conhecimento do facto da soma dos ângulos internos ser de 180° (o modelo funcional) e do conhecimento das propriedades estatísticas dos ângulos observados (o modelo estocástico). Existem diversas formas de obter a representação funcional de um modelo matemático.

- *pré-análise*

A **pré-análise** consiste na especificação da incerteza a exigir às observações, com base na incerteza desejada para os parâmetros desconhecidos e no modelo matemático formulado. Este processo é também conhecido por **análise de covariância**. A pré-análise é feita através da simulação da propagação da incerteza inerente às observações na determinação da incerteza dos parâmetros. Desta forma, é possível, durante o planeamento do procedimento de medição, prever a incerteza dos parâmetros pretendidos e comparar essa incerteza com a desejada. Se essa incerteza for superior (ou muito inferior) à pretendida, o planeamento pode ser alterado antes da realização das medições, minimizando, desta forma, os custos do projecto.

- *medição e pré-processamento*

A informação proveniente da pré-análise é fundamental na concepção das especificações a ter em consideração na **aquisição de dados**. Aquela informação contém, por exemplo, o tipo de observáveis a usar e a incerteza com que estas deverão ser obtidas. Estes dois factores são primordiais na escolha da instrumentação a utilizar e do número de vezes que uma dada medição deverá ser repetida, para satisfazer os requisitos da pré-análise.

O pré-processamento dos dados obtidos visa a correcção de erros sistemáticos e a eliminação de erros grosseiros. Se durante o pré-processamento se verificar que as observações não cumprem os critérios de incerteza, torna-se necessário proceder a novas medições.

Qualquer processo de medição de uma mensuranda está sujeito a erros, devido a factores humanos ("*errare humanum est*"), imperfeições instrumentais e instabilidades do meio ambiente. Os erros classificam-se em 3 categorias principais: grosseiros, sistemáticos e aleatórios.

Os **erros grosseiros** resultam de enganos devidos, essencialmente, à falta de cuidado ou confusão do observador. Como exemplos, pode-se referir leituras erradas de escalas, pontarias erradas (que resultam em valores errados para as observações correspondentes) e registos errados nas folhas de observação. A generalidade dos erros grosseiros pode ser evitada através de verificações de controlo simples, durante a aquisição de dados (efectuar leituras múltiplas e testar a consistência das observações, verificar se a soma dos ângulos internos de um triângulo é próxima dos 180°, etc.). As medições afectadas por erros grosseiros deverão ser omitidas ou corrigidas.

Os **erros sistemáticos** seguem uma lei física (ou podem ser expressos por uma função matemática) e afectam as observações segundo um determinado padrão. Os erros sistemáticos podem ter origem no meio ambiente, instrumentação, observador, ou resultar de múltiplos factores. O meio ambiente, nomeadamente os diversos parâmetros meteorológicos (temperatura, humidade e pressão atmosférica), afecta praticamente todo o tipo de observações. Os principais causas de natureza instrumental relacionam-se com a construção imperfeita dos instrumentos e/ou a sua inadequada calibração. A medição de distâncias com uma fita métrica mal calibrada, por exemplo, resultará num erro sistemático (uma distância final que será maior/menor do que a real se a fita métrica for mais curta/longa, por exemplo). O efeito de um erro sistemático traduz-se numa inconsistência entre as observações e o modelo funcional proposto (neste sentido, as limitações no modelo funcional acabam por resultar em erros sistemáticos).

Existem situações que permitem a eliminação ou redução dos erros sistemáticos. Como exemplo, a magnitude de um erro do tipo instrumental pode ser determinada através da calibração do instrumento, o que permitirá uma correcção das observações. Uma outra forma de redução dos erros sistemáticos consiste em adoptar métodos de observação que atenuem a influência dos erros sistemáticos (muito comum em Topografia). Quando os erros sistemáticos são resultado da influência do meio ambiente, a sua eliminação é quase sempre impossível (por exemplo, a influência da refracção atmosférica na medição de ângulos verticais ou a influência da troposfera na propagação de sinais rádio). Nestes casos, a atenuação dos erros faz-se modificando a estrutura matemática do modelo funcional, introduzindo parâmetros desconhecidos adicionais que venham absorver a sua influência, por exemplo. Estes parâmetros adicionais são chamados **parâmetros inúteis** (designação em literatura inglesa: *nuissance parameters*).

Os **erros aleatórios**, tal como o nome indica, são erros que não exibem qualquer relação funcional baseada num sistema determinístico. Para medições repetidas de uma certa quantidade, estes erros distribuem-se aleatoriamente em torno da média (também conhecido como "ruído branco").

- *ajustamento e solução*

Caso as observações cumpram os critérios estabelecidos, estas são introduzidas no modelo matemático, que conduzirá à **solução**.

- *realização de testes estatísticos*

A verificação da qualidade dos resultados é uma fase importante, que tem como objectivo decidir sobre a aceitação da solução, com base nos critérios de incerteza estabelecidos.

As inconsistências causadas pela variação das observações, que foram encontradas e removidas durante o processo de ajustamento, são usadas para calcular parâmetros estatísticos que, por sua vez, serão utilizados no teste do modelo estocástico e/ou no cálculo de parâmetros específicos de incerteza, como intervalos de confiança e elipses de erro, por exemplo.

- *apresentação e representação gráfica dos resultados*

Caso os resultados obtidos satisfaçam os objectivos definidos, a fase final consiste na **apresentação** e, se aplicável, na **representação gráfica** dos resultados obtidos.

A Figura 1 apresenta a metodologia de análise de dados.

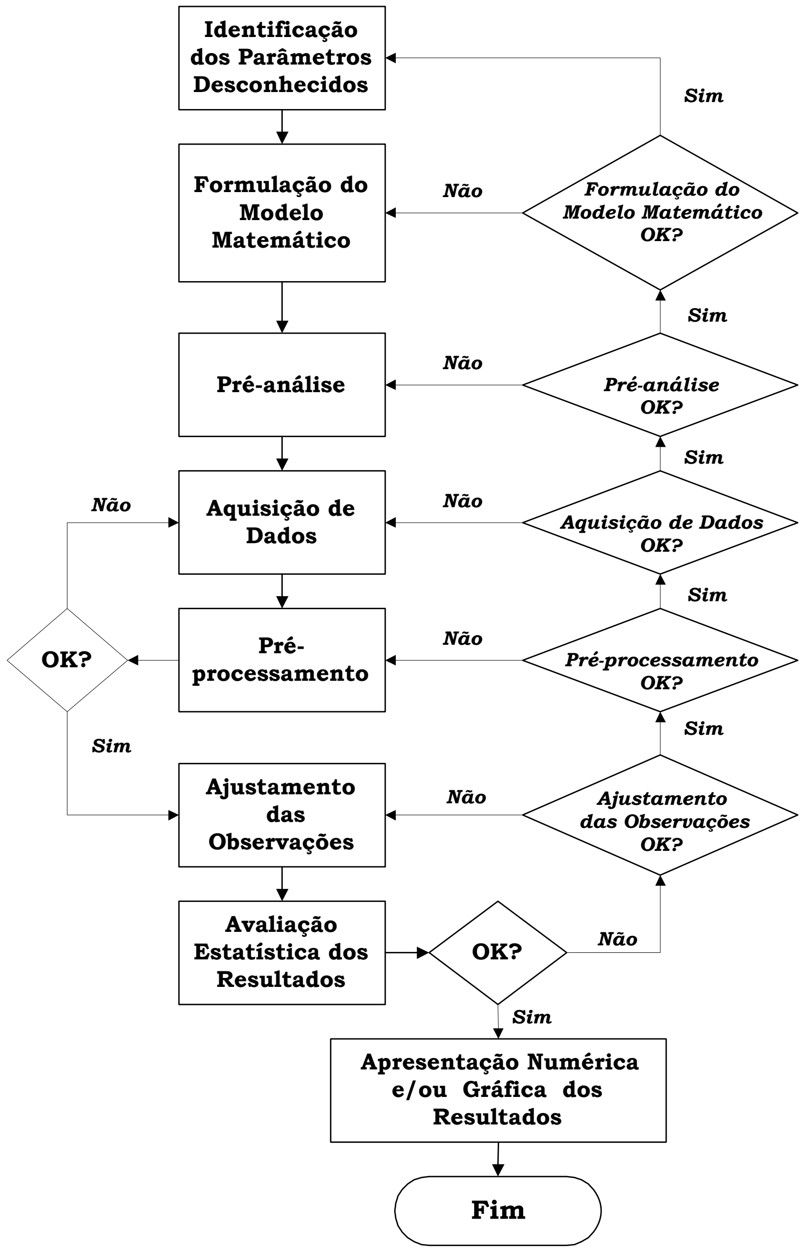

Figura 1 - Metodologia no tratamento e análise de observações em Ciências Geográficas.

2. AJUSTAMENTO POR MÍNIMOS QUADRADOS

Em qualquer tipo de problema, o modelo matemático que traduz a relação entre as observações e os parâmetros desconhecidos (ou apenas entre as observações) é determinado, de forma única, por um número mínimo de medições e/ou constantes que sejam consistentes com o modelo. No entanto, dado que uma medição apresenta sempre um erro associado, em Ciências Geográficas efectua-se um número de medições superior ao necessário. Chama-se **redundância** ou **número de graus de liberdade**, *df*, à diferença entre o número total de observações, *n*, e o número mínimo de observações necessário à determinação de forma única do modelo, n_0:

$$\boxed{\mathrm{df} = \mathrm{n} - \mathrm{n}_0}. \qquad (2.1)$$

As observações redundantes provocam uma inconsistência aparente do modelo matemático: devido aos erros de observação, não existe uma solução única para o problema, ou seja, o número total de observações não se ajusta de forma exacta ao modelo matemático estabelecido. A resolução desta inconsistência aparente passa pela substituição dos valores observados ℓ_i por outro conjunto de observações "corrigidas", que iremos denominar por **observações ajustadas**, $\hat{\ell}_i$: uma vez ajustadas, qualquer subconjunto de observações (as estritamente necessárias) conduzirá à mesma solução.

Uma observação ajustada obtém-se do valor observado adicionando uma correcção v_i, denominada **resíduo**. Em linguagem matricial:

$$\begin{bmatrix} \hat{\ell}_1 \\ \hat{\ell}_2 \\ \hat{\ell}_3 \\ \vdots \\ \hat{\ell}_n \end{bmatrix} = \begin{bmatrix} \ell_1 \\ \ell_2 \\ \ell_3 \\ \vdots \\ \ell_n \end{bmatrix} + \begin{bmatrix} v_1 \\ v_2 \\ v_3 \\ \vdots \\ v_n \end{bmatrix}, \qquad (2.2)$$

ou, de forma condensada:

$$\boxed{\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \boldsymbol{v}}. \qquad (2.3)$$

Denomina-se **ajustamento** à operação de encontrar um conjunto de estimativas $\hat{\ell}_i$ que satisfaça o modelo matemático, de acordo com um dado critério. Esta definição sugere que existem vários critérios que podem conduzir a observações ajustadas; no entanto, existe obviamente vantagem em optar por um critério que nos conduza a uma solução optimizada. No domínio das Ciências Geográficas, e para a generalidade dos casos, é o ajustamento segundo o critério dos mínimos quadrados que satisfaz essa condição.

O **critério dos mínimos quadrados** diz que a soma dos quadrados dos resíduos pesados das observações deverá ser mínimo, ou seja, pretende-se minimizar a função seguinte:

$$\boxed{\phi = \boldsymbol{v}^{\mathrm{T}} \mathbf{P}_{\ell} \boldsymbol{v}}\,, \tag{2.4}$$

onde $\mathbf{P}_{l}$ designa a matriz de pesos das observações.

O conceito de peso está directamente relacionado com o de variância e pode, nesse sentido, ser considerado como uma nova medida de incerteza. A necessidade da introdução deste conceito surge, como veremos, como uma solução de um problema frequente na aplicação do método dos mínimos quadrados, que é o desconhecimento do valor absoluto dos elementos da matriz de variâncias-covariâncias das observações. Para resolver este problema, aplicamos aos elementos da matriz de variâncias-covariâncias um factor de escala arbitrário, denominado **variância *a priori*** (também conhecido como **variância de referência**, **variância (*a priori*) da unidade de peso** ou **factor de variância**).

Comecemos por definir **peso** de uma observação, p_i, à quantidade que é inversamente proporcional à variância dessa mesma observação, σ_i^2:

$$\boxed{p_i = k \frac{1}{\sigma_i^2}}\,, \tag{2.5}$$

onde k é a constante de proporcionalidade. Se consideremos uma observação imaginária com peso unitário, e cuja variância designaremos por σ_0^2, tem-se:

$$1 = \frac{k}{\sigma_0^2} \Leftrightarrow k = \sigma_0^2 .$$

Substituindo em (2.5), obtemos:

$$p_i = \frac{\sigma_0^2}{\sigma_i^2} ,$$

ou ainda,

$$p_1 \cdot \sigma_1^2 = p_2 \cdot \sigma_2^2 = \cdots = k = \sigma_0^2 ,$$

isto é, a constante de proporcionalidade é a variância *a priori* atrás referida.

Da relação anterior torna-se evidente que, para observações não correlacionadas, podemos construir a matriz dos pesos referentes a um conjunto de observações a partir da matriz de variâncias-covariâncias:

$$\mathbf{P}_{\ell} = k\ \mathbf{C}_{\ell}^{-1} .$$

Usando o valor da variância *a priori* como constante de proporcionalidade, a relação anterior pode ser escrita sob a forma

$$\boxed{\mathbf{P}_{\ell} = \sigma_0^2 \mathbf{C}_{\ell}^{-1}} . \tag{2.6}$$

A relação anterior é válida também para os casos em que as observações são correlacionadas. De um modo geral, assume-se para variância *a priori* o valor 1; neste caso particular, a matriz dos pesos é a matriz inversa da matriz de variâncias-covariâncias das observações.

Como já foi referido, existem situações em que os modelos matemáticos traduzem uma relação entre as observações e os parâmetros desconhecidos, enquanto outras podem ser descritas por uma relação existente apenas entre as observações. Designam-se por **equações de observação** as equações que relacionam observações e parâmetros e por **equações de condição** as que relacionam entre si apenas observações ou observações e parâmetros de forma implícita. Esta separação é útil na descrição das diversas técnicas de ajustamento.

Existem várias técnicas de ajustamento por mínimos quadrados, que conduzem a resultados idênticos quando aplicadas ao mesmo problema.

No **ajustamento paramétrico**, também designado por **ajustamento de equações de observação**, **ajustamento de observações indirectas**, ou **ajustamento de parâmetros**:

- as equações que traduzem matematicamente o problema incluem observações e parâmetros, ou seja, são equações de observação;
- o número de equações, m, é igual ao número de observações, n;
- cada equação contém apenas uma observação, com coeficiente unitário, e as quantidades observadas são expressas em função dos parâmetros desconhecidos.

No **ajustamento condicional**, também denominado de **ajustamento (apenas) de observações**, ou **ajustamento de equações de condição**:

- as equações que traduzem matematicamente o problema não contêm parâmetros, ou seja, são equações de condição;
- o número de equações independentes entre as n observações, m, é igual ao número de observações redundantes, df.

Finalmente, no **ajustamento combinado**:

- as equações que traduzem matematicamente o problema incluem observações e parâmetros, mas existe uma relação implícita entre aquelas quantidades, contrariamente ao caso do ajustamento paramétrico;
- o número de equações, m, é a soma do número de observações redundantes, df, com o número de parâmetros desconhecidos independentes, u.

De notar que os ajustamentos paramétrico e condicional são casos particulares de um ajustamento combinado: se o número de parâmetros desconhecidos for nulo, o ajustamento combinado reduz-se a um ajustamento condicional; se o número de parâmetros desconhecidos for igual ao número mínimo de observações, o ajustamento combinado reduz-se a um ajustamento paramétrico. Em resumo:

- Ajustamento combinado: $m = u + df = u + (n - n_0)$;

  $df \leq m \leq n$;

  $0 \leq u \leq n_0$.

- Ajustamento condicional: $u = 0 \Rightarrow m = df$ (limite inferior).
- Ajustamento paramétrico: $u = n_0 \Rightarrow m = n$ (limite superior).

Ajustamento Paramétrico Linear

O modelo matemático para o ajustamento paramétrico é do tipo:

$$\boldsymbol{\ell} = \mathbf{f}(\boldsymbol{x}) .$$

Na sua versão linear, tem-se

$$\boldsymbol{\ell} = \mathbf{A}\boldsymbol{x} + \mathbf{c} , \tag{2.7}$$

onde **c** é um vector de constantes.

Para a dedução das equações para o ajustamento paramétrico, podemos escrever:

$$\mathbf{A}\hat{\boldsymbol{x}} = \hat{\boldsymbol{\ell}} \tag{2.8}$$

Esta equação traduz uma situação ideal, envolvendo observações já ajustadas, que incluem ainda o vector das constantes.

O vector de observações ajustadas pode ser escrito em função das observações e dos resíduos, usando a Equação (2.3),

$$\mathbf{A}\hat{\boldsymbol{x}} - (\boldsymbol{\ell} + \boldsymbol{v}) = \mathbf{0} , \tag{2.9}$$

pelo que

$$\boxed{\boldsymbol{v} = \mathbf{A}\hat{\boldsymbol{x}} - \boldsymbol{\ell}} , \tag{2.10}$$

onde

$\mathbf{A}_{\mathrm{nxu}}$	primeira matriz de configuração (matriz dos coeficientes dos parâmetros)
$\boldsymbol{x}_{\mathrm{u}\times 1}$	vector de parâmetros desconhecidos
$\boldsymbol{\ell}_{\mathrm{n}\times 1}$	vector das observações
$\boldsymbol{v}_{\mathrm{n}\times 1}$	vector de resíduos.

Pretende-se obter estimativas de $\boldsymbol{x}$ que minimizem a forma quadrática dos resíduos pesados

$$\phi = \boldsymbol{v}^{\mathrm{T}} \mathbf{P}_{\ell} \boldsymbol{v} . \tag{2.4}$$

Substituindo a Equação (2.10) na Equação (2.4), vem:

$$\phi = \left(\mathbf{A}\hat{\boldsymbol{x}} - \boldsymbol{\ell}\right)^{\mathrm{T}} \mathbf{P}_{\ell} \left(\mathbf{A}\hat{\boldsymbol{x}} - \boldsymbol{\ell}\right)$$

$$\phi = \left(\left(\mathbf{A}\hat{\boldsymbol{x}}\right)^{\mathrm{T}} - \boldsymbol{\ell}^{\mathrm{T}}\right)\left(\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} - \mathbf{P}_{\ell}\boldsymbol{\ell}\right)$$

$$\phi = \hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} - \boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} - \hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell} + \boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell} \qquad (2.11)$$

Mas $\mathbf{P}_{\ell} = \sigma_0^2\mathbf{C}_{\ell}^{-1}$ e, dado que $\mathbf{C}_{\ell}^{-1}$ é simétrica, $\mathbf{P}_{\ell}^{-1}$ é também simétrica, logo:

$$\mathbf{P}_{\ell}^{\mathrm{T}} = \mathbf{P}_{\ell}\,.$$

Como $\boldsymbol{v}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{v}$ é um escalar, $\hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell}$ e $\boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}}$ serão necessariamente escalares. Assim, podemos escrever:

$$\left(\boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}}\right)^{\mathrm{T}} = \hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell}$$

ou

$$\boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} = \hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell}\,. \qquad (2.12)$$

Substituindo esta relação na Equação (2.11), vem:

$$\phi = \hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} - 2\boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\hat{\boldsymbol{x}} + \boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell}\,.$$

Pretende-se minimizar esta função em relação a $\hat{\boldsymbol{x}}$, ou seja:

$$\frac{\partial}{\partial\hat{\mathbf{x}}}\left(\boldsymbol{v}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{v}\right) = \mathbf{0}\,.$$

Usando as relações (ver Apêndice 1)

$$\frac{\partial}{\partial\hat{x}}\left(\mathbf{A}\hat{x}\right) = \mathbf{A}\,,$$

$$\frac{\partial}{\partial\hat{\mathbf{x}}}\left(\hat{\boldsymbol{x}}^{\mathrm{T}}\mathrm{A}\hat{\boldsymbol{x}}\right) = 2\hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}$$

e a Equação (2.12), tem-se

$$\frac{\partial}{\partial\hat{\mathbf{x}}}\left(\boldsymbol{v}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{v}\right) = 2\hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A} - 2\boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A} = \mathbf{0}$$

$$\hat{\boldsymbol{x}}^{\mathrm{T}}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\mathbf{A} = \boldsymbol{\ell}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\mathbf{A}\,.$$

Transpondo ambos os membros desta equação obtemos

$$\boxed{\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\mathbf{A}\right)\hat{x} = \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\boldsymbol{\ell}}\,. \tag{2.13}$$

Esta expressão constitui o **sistema de equações normais**. Abreviadamente, pode-se escrever:

$$\mathbf{N}\hat{\boldsymbol{x}} = \mathbf{U}\,,$$

onde $\mathbf{N} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\mathbf{A}\right)$ designa a **matriz dos coeficientes das equações normais** e $\mathbf{U} = \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\boldsymbol{\ell}$ o **vector dos termos absolutos das equações normais**. Desta equação obtemos para o vector das estimativas de $\hat{\boldsymbol{x}}$:

$$\hat{\boldsymbol{x}} = \mathbf{N}^{-1}\,\mathbf{U}\,, \tag{2.14}$$

ou, de forma explícita,

$$\boxed{\hat{\boldsymbol{x}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\boldsymbol{\ell}}\,. \tag{2.15}$$

De notar que N tem sempre inversa, dado que é uma matriz simétrica definida positiva.

Uma vez determinado o vector de parâmetros desconhecidos, este pode ser substituído na Equação (2.10), o que permite determinar o vector de resíduos ajustados das observações, a partir do qual determinamos o vector de observações ajustadas:

$$\boxed{\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \hat{\boldsymbol{v}}}\,. \tag{2.16}$$

A estimativa $\hat{\boldsymbol{x}}$ é independente da escolha de σ_0^2, como se pode demonstrar. Consideremos uma nova matriz

$$\mathbf{P}'_{\boldsymbol{\ell}} = \gamma\mathbf{P}_{\boldsymbol{\ell}} \tag{2.17}$$

onde γ é uma constante arbitrária. Substituindo na Equação (2.15), obtemos uma nova estimativa para os parâmetros desconhecidos, $\hat{\boldsymbol{x}}'$:

$$\hat{x}' = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}'_{\ell}\,\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}'_{\ell}\,\boldsymbol{\ell} = \left(\mathbf{A}^{\mathrm{T}}\gamma\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\gamma\mathbf{P}_{\ell}\boldsymbol{\ell}$$

$$\hat{x}' = \frac{1}{\gamma}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\gamma\mathbf{P}_{\ell}\boldsymbol{\ell} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{\ell} = \hat{x}\,.$$

Este facto permite-nos afirmar que para obtermos o vector de parâmetros desconhecidos (bem como a estimativa da matriz de variâncias-covariâncias correspondente, como veremos), poderemos assumir qualquer peso relativo para as observações, ou seja, um σ_0^2 arbitrário.

Matrizes de variâncias-covariâncias

A **matriz de variâncias-covariâncias dos parâmetros ajustados** deduz-se facilmente, aplicando a lei de propagação de variâncias-covariâncias. Tem-se:

$$\mathbf{C}_{\hat{x}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}}\,, \tag{2.18}$$

sendo

$$\mathbf{J} = \frac{\partial\hat{\mathbf{x}}}{\partial\boldsymbol{\ell}} = \mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\,. \tag{2.19}$$

Dado que

$$\mathbf{P}_{\ell} = \sigma_o^2\mathbf{C}_{\ell}^{-1}\,,$$

então

$$\mathbf{C}_{\ell} = \sigma_o^2\mathbf{P}_{\ell}^{-1}\,.$$

Substituindo na Equação (2.18), obtemos

$$\mathbf{C}_{\hat{x}} = \left(\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\right)\sigma_0^2\mathbf{P}_{\ell}^{-1}\left(\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\right)^{\mathrm{T}}\,. \tag{2.20}$$

Dado que $\mathbf{P}_{\ell}$ e **N** são matrizes simétricas, são válidas as seguintes relações:

$$\mathbf{P}_{\ell}^{\mathrm{T}} = \mathbf{P}_{\ell}$$

$$\mathbf{N}^{\mathrm{T}} = \mathbf{N}$$

$$\left(\mathbf{N}^{-1}\right)^{\mathrm{T}} = \mathbf{N}^{-1}\,.$$

Substituindo na Equação (2.20), vem:

$$\mathbf{C}_{\hat{x}} = \sigma_0^2 \mathbf{N}^{-1} \mathbf{A}^{\mathrm{T}} \cancel{\mathbf{P}_\ell} \cancel{\mathbf{P}_\ell^{-1}} \mathbf{P}_\ell \mathbf{A} \mathbf{N}^{-1} = \sigma_0^2 \mathbf{N}^{-1} \underbrace{\left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_\ell \mathbf{A}\right)}_{\mathbf{N}} \mathbf{N}^{-1}$$

$$\mathbf{C}_{\hat{x}} = \sigma_0^2 \cancel{\mathbf{N}^{-1}} \cancel{\mathbf{N}} \mathbf{N}^{-1}$$

$$\mathbf{C}_{\hat{x}} = \sigma_0^2 \mathbf{N}^{-1},$$

ou seja,

$$\boxed{\mathbf{C}_{\hat{x}} = \sigma_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_\ell \mathbf{A}\right)^{-1}}. \tag{2.21}$$

Fazendo $\mathbf{P}_\ell = \sigma_o^2 \mathbf{C}_\ell^{-1}$ nesta equação, obtemos:

$$\mathbf{C}_{\hat{x}} = \frac{1}{\sigma_0^2} \sigma_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{C}_\ell^{-1} \mathbf{A}\right)^{-1}$$

$$\boxed{\mathbf{C}_{\hat{x}} = \left(\mathbf{A}^{\mathrm{T}} \mathbf{C}_\ell^{-1} \mathbf{A}\right)^{-1}} \tag{2.22}$$

Conclui-se, assim, que se as variâncias (e covariâncias) absolutas das observações forem conhecidas, $\mathbf{C}_{\hat{x}}$ não depende da escolha de σ_0^2. Por vezes, no entanto, apenas as variâncias (e covariâncias) relativas das observações são conhecidas, pelo que se deve usar a matriz de pesos e não a matriz de variâncias-covariâncias no cálculo da matriz de variâncias-covariâncias dos parâmetros ajustados. Embora se desconheça à partida a variância *a priori* σ_0^2, esta poderá ser estimada a partir dos resíduos ajustados, sendo o seu valor dado por:

$$\boxed{\hat{\sigma}_0^2 = \frac{\hat{v}^{\mathrm{T}} \mathbf{P}_\ell \hat{v}}{\mathrm{df}}}. \tag{2.23}$$

O valor estimado para a variância *a priori*, denominada **variância *a posteriori*** (ou **variância de referência *a posteriori***), pode agora ser usado para determinar a **estimativa da matriz de variâncias-covariâncias de $\hat{x}$**:

$$\boxed{\hat{\mathbf{C}}_{\hat{x}} = \hat{\sigma}_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_\ell \mathbf{A}\right)^{-1}}. \tag{2.24}$$

Tal como no caso do vector de parâmetros ajustados, a estimativa da matriz de variâncias-covariâncias de $\hat{\boldsymbol{x}}$ não depende da escolha da variância *a priori*. Com efeito, consideremos uma nova matriz de pesos $\mathbf{P}'_{\boldsymbol{\ell}} = \gamma \mathbf{P}_{\boldsymbol{\ell}}$, onde γ é uma constante arbitrária. Substituindo a Equação (2.23) na Equação (2.24), e introduzindo a nova matriz de pesos, obtemos uma nova estimativa $\hat{\mathbf{C}}'_{\hat{x}}$:

$$\hat{\mathbf{C}}'_{\hat{x}} = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}'_{\boldsymbol{\ell}} \hat{\boldsymbol{v}}}{\mathrm{df}} \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}'_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1} = \frac{\gamma \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \hat{\boldsymbol{v}}}{\mathrm{df}} \frac{1}{\gamma} \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1}$$

ou seja,

$$\hat{\mathbf{C}}'_{\hat{x}} = \hat{\sigma}_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1} = \hat{\mathbf{C}}_{\hat{x}} \,.$$

A Tabela 1 sintetiza o conjunto de equações do ajustamento paramétrico linear.

Tabela 1 - Equações para o ajustamento paramétrico linear.

$$\hat{\boldsymbol{x}} = \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1} \mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \boldsymbol{\ell}$$

$$\hat{\boldsymbol{v}} = \mathbf{A} \hat{\boldsymbol{x}} - \boldsymbol{\ell}$$

$$\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \hat{\boldsymbol{v}}$$

$$\mathbf{C}_{\hat{x}} = \sigma_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1}$$

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \hat{\boldsymbol{v}}}{\mathrm{df}}$$

$$\hat{\mathbf{C}}_{\hat{x}} = \hat{\sigma}_0^2 \left(\mathbf{A}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \mathbf{A} \right)^{-1}$$

Ajustamento Condicional Linear

O modelo matemático para um ajustamento condicional é do tipo

$$\mathbf{f}(\hat{\boldsymbol{\ell}}) = \mathbf{0}.$$

A versão linear correspondente tem a forma

$$\mathbf{B}\boldsymbol{\ell} = \mathbf{c}, \tag{2.25}$$

equação que representa um conjunto de m condições linearmente <u>independentes</u> entre n observações, sendo $\mathbf{c}$ um vector de constantes. Para obter as equações para o ajustamento condicional, a Equação (2.25) pode ser representada sob a forma

$$\mathbf{B}\hat{\boldsymbol{\ell}} = \mathbf{c}. \tag{2.26}$$

Introduzindo a Equação (2.3), podemos substituir o vector de observações ajustadas pelo vector das observações, obtendo-se

$$\mathbf{B}\left(\boldsymbol{\ell} + \hat{\boldsymbol{v}}\right) - \mathbf{c} = \mathbf{0}, \tag{2.27}$$

ou, de forma equivalente,

$$\boxed{\mathbf{B}\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}}, \tag{2.28}$$

onde

$$\boldsymbol{w} = \mathbf{B}\boldsymbol{\ell} - \mathbf{c}. \tag{2.29}$$

Nas equações anteriores, tem-se:

- $\mathbf{B}_{mxn}$ segunda matriz de configuração (matriz dos coeficientes dos resíduos ou matriz dos coeficientes das observações)
- $\mathbf{c}_{mx1}$ vector de constantes
- $\boldsymbol{\ell}_{n\times1}$ vector de observações
- $\boldsymbol{w}_{mx1}$ vector de fecho ("*misclosure*")
- $\boldsymbol{v}_{nx1}$ vector de resíduos.

Note-se que são respeitadas as características inerentes ao ajustamento condicional:

- não existem parâmetros desconhecidos nas equações de observação (u=0);
- existem df equações de condição entre n observações (m = df, limite inferior).

As incógnitas nas equações de condição definidas pela Equação (2.28) são os resíduos ajustados.

Utilizando o critério dos mínimos quadrados, pretendemos determinar um conjunto de estimativas para $\boldsymbol{v}$ que minimize a soma dos quadrados dos resíduos pesados:

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} .$$

Como em geral $\hat{\boldsymbol{v}}$ não pode ser expresso em função de $\mathbf{B}$ e $\boldsymbol{w}$ (dado que $\mathbf{B}$ não é invertível), vamos introduzir um vector $\boldsymbol{\lambda}$ de df incógnitas, denominado **vector de multiplicadores de Lagrange**:

$$\boldsymbol{\lambda}_{\mathrm{m}\times 1} = 2\mathbf{k}^{\mathrm{T}} .$$

Consideremos uma nova função ϕ', dada por

$$\phi' = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} + 2\mathbf{k}^{\mathrm{T}} \left(\mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} \right) .$$

Como

$$\mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0} ,$$

tem-se

$$\phi' = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} = \phi .$$

A equação $\mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0}$ representa o constrangimento (ou condição) em relação ao qual ϕ terá que ser mínimo.

Dado que

$$\frac{\partial}{\partial \hat{\boldsymbol{v}}} \left(\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} \right) = 2\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} ,$$

a minimização da função ϕ é dada por:

$$\frac{\partial \phi}{\partial \hat{\boldsymbol{v}}} = 2\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} + 2\hat{\mathbf{k}}^{\mathrm{T}} \mathbf{B} = \mathbf{0} ,$$

ou

$$\hat{v}^{\mathrm{T}}\mathbf{P}_{\ell} + \hat{\mathbf{k}}^{\mathrm{T}}\mathbf{B} = \mathbf{0}\,.$$

Transpondo ambos os membros da equação, obtemos:

$$\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} + \mathbf{P}_{\ell}\hat{v} = \mathbf{0}\,,$$

ou seja,

$$\hat{v} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}\,.$$

Substituindo esta equação na Equação (2.28), vem:

$$\mathbf{B}\left(-\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}\right) + w = \mathbf{0}\,,$$

ou

$$\boxed{\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)\hat{\mathbf{k}} = w}\,. \tag{2.30}$$

A Equação (2.30) constitui o **sistema de equações normais para o ajustamento condicional** (note-se a semelhança com o sistema de equações normais para o ajustamento paramétrico).

Nesta fase, o vector de multiplicadores de Lagrange ainda não é conhecido. Introduzindo a matriz auxiliar

$$\mathbf{M} = \mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\,, \tag{2.31}$$

podemos escrever (2.30) sob a forma

$$\mathbf{M}\hat{\mathbf{k}} = w\,. \tag{2.32}$$

A solução do sistema de equações normais anterior em ordem a $\hat{\mathbf{k}}$ é

$$\boxed{\hat{\mathbf{k}} = \mathbf{M}^{-1}w}\,. \tag{2.33}$$

Obtido o vector de multiplicadores de Lagrange, podemos agora calcular o vector de resíduos ajustados

$$\boxed{\hat{v} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}} \tag{2.34}$$

e, finalmente, o vector das observações ajustadas

$$\boxed{\hat{\ell} = \ell + \hat{v}}. \tag{2.35}$$

O conjunto de expressões anteriores pode ser combinado numa expressão única:

$$\hat{\ell} = \ell - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}w\,, \tag{2.36}$$

ou ainda,

$$\boxed{\hat{\ell} = \ell - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{B}\ell - \mathbf{c}\right)}. \tag{2.37}$$

Matrizes de variâncias-covariâncias

A matriz de variâncias-covariâncias para as observações ajustadas obtém-se aplicando a lei de propagação de variâncias-covariâncias.

Tem-se

$$\mathbf{C}_{\hat{\ell}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}}\,,$$

sendo

$$\mathbf{J} = \frac{\partial\hat{\boldsymbol{\ell}}}{\partial\boldsymbol{\ell}} = \mathbf{I} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\,.$$

Substituindo, obtemos:

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{I} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\right)\mathbf{P}_{\ell}^{-1}\left(\mathbf{I} - \mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{I} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\right)\left(\mathbf{P}_{\ell}^{-1} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{P}_{\ell}^{-1} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1} - \cancel{\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}} + \cancel{\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\underbrace{\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}}_{\cancel{\mathbf{M}}}\cancel{\mathbf{M}}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}}\right),$$

ou seja,

$$\boxed{\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{P}_\ell^{-1} - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)}. \tag{2.38}$$

Recordando que $\mathbf{C}_\ell = \sigma_0^2\mathbf{P}_\ell^{-1}$, a Equação (2.38) pode ser escrita sob a forma

$$\boxed{\mathbf{C}_{\hat{\ell}} = \mathbf{C}_\ell - \sigma_0^2\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}}. \tag{2.39}$$

Se definirmos

$$\mathbf{T} = \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}, \tag{2.40}$$

tem-se, em alternativa, a seguinte representação:

$$\boxed{\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{I} - \mathbf{T}\right)\mathbf{P}_\ell^{-1}}. \tag{2.41}$$

Desta relação podemos concluir que a matriz de variâncias-covariâncias das observações ajustadas pode vir a ser não diagonal, mesmo se a matriz de variâncias-covariâncias das observações for diagonal; isto deve-se à introdução das equações de condição, que provocam correlação nas observações ajustadas. Se a matriz dos coeficientes das observações for ortogonal, são válidas as seguintes relações:

$$\mathbf{B}^{\mathrm{T}} = \mathbf{B}^{-1},$$

e

$$\mathbf{B}\mathbf{B}^{\mathrm{T}} = \mathbf{B}^{\mathrm{T}}\mathbf{B} = \mathbf{I}.$$

Nesta condição, e não existindo correlação entre as observações (covariâncias nulas), as matrizes $\mathbf{T}$ e $\mathbf{P}_\ell^{-1}$ serão matrizes diagonais e, para este caso particular, as observações ajustadas serão não correlacionadas. Note-se ainda que, uma vez que os elementos da diagonal principal das observações não ajustadas são positivos e que $\mathbf{T}$ é uma matriz definida positiva, se obtém uma incerteza para as observações ajustadas que é inferior à incerteza das observações não ajustadas.

Tal como no ajustamento paramétrico, podemos obter a variância *a posteriori*, usando a relação (2.23):

$$\boxed{\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \hat{\boldsymbol{v}}}{\mathrm{df}}} \tag{2.23}$$

Uma vez obtida a variância *a posteriori*, e de novo por analogia com o ajustamento paramétrico, podemos obter a estimativa da matriz de variâncias-covariâncias das observações ajustadas:

$$\boxed{\hat{\mathbf{C}}_{\hat{\ell}} = \hat{\sigma}_0^2 \left(\mathbf{I} - \mathbf{T}\right) \mathbf{P}_{\boldsymbol{\ell}}^{-1}}. \tag{2.42}$$

Será também fácil demonstrar que a escolha da matriz de pesos (que mantenha a incerteza relativa das observações, ou seja, para qualquer valor de variância *a priori*) não influencia quer o vector de observações ajustadas quer a estimativa da matriz de variâncias-covariâncias correspondente.

O conjunto de equações a utilizar num ajustamento condicional está resumido na Tabela 2.

Tabela 2 – Equações para o ajustamento condicional linear.

$$\boldsymbol{w} = \mathbf{B}\boldsymbol{\ell} - \mathbf{c}$$

$$\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_{\boldsymbol{\ell}}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1} \boldsymbol{w}$$

$$\hat{\boldsymbol{v}} = -\mathbf{P}_{\boldsymbol{\ell}}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{k}$$

$$\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \hat{\boldsymbol{v}}$$

$$\mathbf{T} = \mathbf{P}_{\boldsymbol{\ell}}^{-1}\left[\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\boldsymbol{\ell}}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\right]$$

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2 \left(\mathbf{I} - \mathbf{T}\right)\mathbf{P}_{\boldsymbol{\ell}}^{-1}$$

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\boldsymbol{\ell}} \hat{\boldsymbol{v}}}{\mathrm{df}}$$

$$\hat{\mathbf{C}}_{\hat{\ell}} = \hat{\sigma}_0^2 \left(\mathbf{I} - \mathbf{T}\right)\mathbf{P}_{\boldsymbol{\ell}}^{-1}$$

Ajustamento Combinado

Em Ciências Geográficas, os modelos matemáticos que descrevem as situações físicas são frequentemente do tipo não linear. Para obter um conjunto de equações que permita fazer um ajustamento de observações para modelos não lineares, o modelo matemático é linearizado, usando expansão em séries de Taylor. Para tal, é necessário partir de uma aproximação inicial dos parâmetros desconhecidos e determinar correcções a estes valores iniciais.

O modelo matemático para o método combinado é do tipo:

$$\mathbf{f}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}\right)=\mathbf{0}\,. \tag{2.43}$$

Este modelo expressa uma relação não linear de m equações entre n observações e u parâmetros desconhecidos, com

$$\mathrm{m}=\mathrm{u}+\mathrm{df}\,.$$

Linearizando em torno de valores aproximados dos parâmetros desconhecidos $\boldsymbol{x}_0$ e do vector de observações $\boldsymbol{\ell}$, vem

$$\mathbf{f}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}\right)\doteq\mathbf{f}\left(\boldsymbol{x}_0,\boldsymbol{\ell}\right)+\left.\frac{\partial\mathbf{f}}{\partial\hat{\mathbf{x}}}\right|_{x_0,\,\ell}\left(\boldsymbol{x}-\boldsymbol{x}_0\right)+\left.\frac{\partial\mathbf{f}}{\partial\boldsymbol{\ell}}\right|_{x_0,\,\ell}\left(\hat{\boldsymbol{\ell}}-\boldsymbol{\ell}\right)=\mathbf{0} \tag{2.44}$$

ou, em representação matricial,

$$\boxed{\mathbf{A}\delta+\mathbf{B}\boldsymbol{v}+\boldsymbol{w}=\mathbf{0}}, \tag{2.45}$$

onde

$\mathbf{A}_{\mathrm{m}\times\mathrm{u}}=\left.\dfrac{\partial\mathbf{f}}{\partial\hat{\mathbf{x}}}\right|_{x_0,\ell}$ primeira matriz de configuração

(matriz de coeficientes dos parâmetros/correcções)

$\mathbf{B}_{\mathrm{m}\times\mathrm{n}}=\left.\dfrac{\partial\mathbf{f}}{\partial\hat{\boldsymbol{\ell}}}\right|_{x_0,\ell}$ segunda matriz de configuração

(matriz de coeficientes das observações/resíduos)

$\delta_{u\times1} = \left(\hat{\boldsymbol{x}} - \boldsymbol{x}_0\right)$ vector de correcções aos parâmetros desconhecidos

$\boldsymbol{v}_{n\times1} = \hat{\boldsymbol{\ell}} - \boldsymbol{\ell}$ vector de resíduos

$\boldsymbol{w}_{m\times1} = \mathbf{f}\left(\boldsymbol{x}_0, \boldsymbol{\ell}\right)$ vector de fecho

Pretendemos minimizar a função

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}}\,.$$

Tal como no método condicional, é necessário introduzir o vector de multiplicadores de Lagrange, pelo que a função a minimizar poderá ser escrita sob a forma

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} + 2\hat{\mathbf{k}}^{\mathrm{T}}\left(\mathbf{A}\hat{\delta} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w}\right).$$

Calculando as derivadas parciais em relação a $\hat{\boldsymbol{v}}$ e a $\hat{\delta}$ obtemos, respectivamente:

$$\frac{\partial\phi}{\partial\hat{\boldsymbol{v}}} = 2\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} + 2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{B} = \mathbf{0} \tag{2.46}$$

e

$$\frac{\partial\phi}{\partial\hat{\delta}} = 2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{A} = \mathbf{0}\,. \tag{2.47}$$

Transpondo ambas as equações (e dividindo por 2), vem:

$$\mathbf{P}_{\ell}\hat{\boldsymbol{v}} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \tag{2.48}$$

$$\mathbf{A}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}\,. \tag{2.49}$$

Necessitamos encontrar uma solução que satisfaça as equações (2.45), (2.48) e (2.49):

$$\begin{cases} \mathbf{A}\hat{\delta} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0} \\ \mathbf{P}_{\ell}\hat{\boldsymbol{v}} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \\ \mathbf{A}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \end{cases}$$

Estas três equações podem ser combinadas numa hipermatriz (ou seja, uma matriz cujos elementos são eles próprios matrizes):

$$\begin{bmatrix} {}_{n\times n}\mathbf{P}_\ell & {}_{n\times m}\mathbf{B}^{\mathrm{T}} & {}_{n\times u}\mathbf{0} \\ {}_{m\times n}\mathbf{B} & {}_{m\times m}\mathbf{0} & {}_{m\times u}\mathbf{A} \\ {}_{u\times n}\mathbf{0} & {}_{u\times m}\mathbf{A}^{\mathrm{T}} & {}_{u\times u}\mathbf{0} \end{bmatrix} \begin{bmatrix} {}_{n\times 1}\hat{\boldsymbol{v}} \\ {}_{m\times 1}\hat{\mathbf{k}} \\ {}_{u\times 1}\hat{\boldsymbol{\delta}} \end{bmatrix} + \begin{bmatrix} {}_{n\times 1}\mathbf{0} \\ {}_{m\times 1}\boldsymbol{w} \\ {}_{u\times 1}\mathbf{0} \end{bmatrix} = \mathbf{0}, \tag{2.50}$$

com as dimensões de cada submatriz (excepcionalmente) indicadas à esquerda. Estabelecendo a seguinte partição por blocos:

$$\left[\begin{array}{c:cc} \mathbf{P}_\ell & \mathbf{B}^{\mathrm{T}} & \mathbf{0} \\ \hdashline \mathbf{B} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\boldsymbol{v}} \\ \hdashline \hat{\mathbf{k}} \\ \hat{\boldsymbol{\delta}} \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline \boldsymbol{w} \\ \mathbf{0} \end{array}\right] = \mathbf{0}, \tag{2.51}$$

tem-se:

$$\begin{aligned} &\mathbf{A}_{11} = \mathbf{P}_\ell, \\ &\mathbf{A}_{12} = \begin{bmatrix} \mathbf{B}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}, \\ &\mathbf{A}_{21} = \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix}, \\ &\mathbf{A}_{22} = \begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}, \\ &\mathbf{X}_1 = \hat{\boldsymbol{v}}, \\ &\mathbf{X}_2 = \begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\boldsymbol{\delta}} \end{bmatrix}, \\ &\mathbf{U}_1 = \mathbf{0}, \\ &\mathbf{U}_2 = \begin{bmatrix} \boldsymbol{w} \\ \mathbf{0} \end{bmatrix}. \end{aligned} \tag{2.52}$$

Usando as regras de multiplicação de matrizes por blocos (particionadas), podemos estabelecer as seguintes relações:

$$\mathbf{A}_{11}\mathbf{X}_1 + \mathbf{A}_{12}\mathbf{X}_2 + \mathbf{U}_1 = \mathbf{0} \tag{2.53}$$

$$\mathbf{A}_{21}\mathbf{X}_1 + \mathbf{A}_{22}\mathbf{X}_2 + \mathbf{U}_2 = \mathbf{0}. \tag{2.54}$$

Admitindo que A_{11} tem inversa, da Equação (2.53) obtemos

$$\mathbf{X}_1 = -\mathbf{A}_{11}^{-1}\left(\mathbf{U}_1 + \mathbf{A}_{12}\mathbf{X}_2\right). \tag{2.55}$$

Substituindo a Equação (2.55) na Equação (2.54),

$$-\mathbf{A}_{21}\mathbf{A}_{11}^{-1}\left(\mathbf{U}_1 + \mathbf{A}_{12}\mathbf{X}_2\right) + \mathbf{A}_{22}\mathbf{X}_2 + \mathbf{U}_2 = \mathbf{0}, \tag{2.56}$$

e pondo $\mathbf{X}_2$ em evidência, obtemos:

$$\boxed{\left(\mathbf{A}_{22} - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{A}_{12}\right)\mathbf{X}_2 + \left(\mathbf{U}_2 - \mathbf{A}_{21}\mathbf{A}_{11}^{-1}\mathbf{U}_1\right) = \mathbf{0}}. \tag{2.57}$$

Substituindo as Equações (2.52) na Equação (2.57), obtemos:

$$\left[\begin{bmatrix} \mathbf{0} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{P}_{\ell}^{-1} \begin{bmatrix} \mathbf{B}^{\mathrm{T}} & \mathbf{0} \end{bmatrix}\right] \begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\delta} \end{bmatrix} + \left[\begin{bmatrix} w \\ \mathbf{0} \end{bmatrix} - \begin{bmatrix} \mathbf{B} \\ \mathbf{0} \end{bmatrix} \mathbf{P}_{\ell}^{-1} \begin{bmatrix} \mathbf{0} \end{bmatrix}\right] = \mathbf{0} \tag{2.58}$$

ou seja:

$$\begin{bmatrix} -\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\mathbf{k}} \\ \hat{\delta} \end{bmatrix} + \begin{bmatrix} w \\ \mathbf{0} \end{bmatrix} = \mathbf{0}. \tag{2.59}$$

Para determinar $\hat{\delta}$ particiona-se a hipermatriz de modo a que $\mathbf{X}_1 = \hat{\mathbf{k}}$:

$$\left[\begin{array}{c:c} -\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \hdashline \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}} \\ \hdashline \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w \\ \hdashline \mathbf{0} \end{array}\right] = \mathbf{0} \tag{2.60}$$

O novo conjunto de submatrizes é dado por:

$$\begin{aligned} \mathbf{A}_{11} &= -\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}, \\ \mathbf{A}_{12} &= \mathbf{A}, \\ \mathbf{A}_{21} &= \mathbf{A}^{\mathrm{T}}, \\ \mathbf{A}_{22} &= \mathbf{0}, \\ \mathbf{X}_1 &= \hat{\mathbf{k}}, \\ \mathbf{X}_2 &= \hat{\delta}, \\ \mathbf{U}_1 &= w, \\ \mathbf{U}_2 &= \mathbf{0} \end{aligned} \tag{2.61}$$

Usando a relação (2.57), temos:

$$\left[\mathbf{0}-\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right]\hat{\delta}+\left[\mathbf{0}-\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{w}\right]=\mathbf{0}, \tag{2.62}$$

donde

$$\boxed{\hat{\delta}=-\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{w}}. \tag{2.63}$$

A estimativa para $\hat{\mathbf{k}}$ obtém-se da primeira equação de (2.60):

$$-\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)\hat{\mathbf{k}}+\mathbf{A}\hat{\delta}+\boldsymbol{w}=\mathbf{0}, \tag{2.64}$$

pelo que

$$\boxed{\hat{\mathbf{k}}=\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}\hat{\delta}+\boldsymbol{w}\right)}. \tag{2.65}$$

Finalmente, para obtermos uma estimativa para o vector de resíduos, recorremos ao sistema de equações (2.50). Para a primeira equação desse sistema, vem:

$$\mathbf{P}_{\ell}\hat{\boldsymbol{v}}+\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}=\mathbf{0} \tag{2.66}$$

donde

$$\boxed{\hat{\boldsymbol{v}}=-\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}}. \tag{2.67}$$

As soluções $\hat{\delta}$ e $\hat{\boldsymbol{v}}$ devem ser adicionadas às aproximações iniciais $\boldsymbol{x}_0$ e $\boldsymbol{\ell}$, de forma a obter as estimativas do vectores de solução e de observação ajustadas:

$$\boxed{\hat{\boldsymbol{x}}=\boldsymbol{x}_0+\hat{\delta}} \tag{2.68}$$

$$\boxed{\hat{\boldsymbol{\ell}}=\boldsymbol{\ell}+\hat{\boldsymbol{v}}}. \tag{2.69}$$

Matrizes de variâncias-covariâncias

As expressões para as matrizes de variâncias-covariâncias dos parâmetros desconhecidos e das observações ajustadas são obtidas, como é usual, usando a lei de propagação de variâncias-covariâncias.

Para obter a **matriz de variâncias-covariâncias para** $\hat{\delta}$, comecemos por recordar que:

$$w = \mathbf{f}\left(x_0, \ell\right)$$

$$\mathbf{C}_\ell = \sigma_0^2 \mathbf{P}_\ell^{-1}$$

$$\mathbf{M} = \mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}, \tag{2.31}$$

e

$$\frac{\partial \mathbf{f}\left(x_0, \ell\right)}{\partial \hat{\ell}} = \frac{\partial \mathbf{f}\left(x_0, \ell\right)}{\partial \ell} = \mathbf{B}$$

Tem-se

$$\mathbf{C}_{\hat{\delta}} = \mathbf{J}\mathbf{C}_\ell \mathbf{J}^{\mathrm{T}}, \tag{2.70}$$

sendo

$$\mathbf{J} = \frac{\partial \hat{\delta}}{\partial \ell} = -\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}.$$

Substituindo na Equação (2.70), tem-se:

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\left(\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\right)\mathbf{P}_\ell^{-1}\left(\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\right)\right)$$

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\underbrace{\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}}_{\cancel{\mathbf{M}}}\cancel{\mathbf{M}}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\right)$$

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\cancel{\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)}\cancel{\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}}\right)$$

$$\boxed{\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}}. \tag{2.71}$$

Dado que

$$\hat{x} = x_0 + \hat{\delta},$$

onde x_0 é constante, tem-se:

$$\boxed{\mathbf{C}_{\hat{x}} = \mathbf{C}_{\hat{\delta}}}. \tag{2.72}$$

A **matriz de variâncias-covariâncias para $\hat{\ell}$** pode ser derivada diretamente a partir da forma explícita para aquele vector. Tem-se:

$$\hat{\ell} = \ell - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\left(\mathbf{A}\hat{\delta} + w\right)$$

$$\hat{\ell} = \ell + \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}w - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}w \tag{2.73}$$

Tem-se

$$\mathbf{C}_{\hat{\ell}} = \mathbf{J}\mathbf{C}_\ell\mathbf{J}^{\mathrm{T}},$$

sendo

$$\mathbf{J} = \frac{\partial\hat{\ell}}{\partial\ell} = \mathbf{I} + \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}.$$

Substituindo na Equação (2.73), com $\mathbf{C}_\ell = \sigma_0^2\mathbf{P}_\ell^{-1}$, e introduzindo uma matriz auxiliar **N** (matriz de equações normais para o ajustamento combinado**)**

$$\mathbf{N} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1} \tag{2.74}$$

temos:

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left[\left(\mathbf{P}_\ell^{-1} + \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1} - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)\right.$$
$$\left.\cdot\left(\mathbf{I} + \mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1} - \mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)\right].$$

Desenvolvendo e simplificando, obtém-se

$$\boxed{\mathbf{C}_{\hat{\ell}} = \sigma_0^2\mathbf{P}_\ell^{-1} - \sigma_0^2\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\left(\mathbf{B}\mathbf{P}_\ell^{-1} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)}. \tag{2.75}$$

Usando um procedimento semelhante, prova-se que a matriz de variâncias-covariâncias para os resíduos ajustados é dada pela expressão

$$\boxed{\mathbf{C}_{\hat{v}} = \sigma_0^2\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\left(\mathbf{B}\mathbf{P}_\ell^{-1} - \mathbf{A}\mathbf{N}^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)}, \tag{2.76}$$

pelo que se pode estabelecer a seguinte relação:

$$\boxed{\mathbf{C}_{\hat{\ell}} = \mathbf{C}_\ell - \mathbf{C}_{\hat{v}}}. \tag{2.77}$$

Tal como nos casos dos ajustamentos de funções lineares, podemos estimar a variância *a posteriori* a partir dos resíduos ajustados,

$$\boxed{\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}}}{\mathrm{df}}}, \tag{2.23}$$

e obter estimativas para as matrizes de variâncias-covariâncias dos parâmetros ajustados e das observações ajustadas.

Em todo este processo de obtenção de estimativas para os parâmetros desconhecidos e observações ajustadas utilizámos quantidades (as matrizes de configuração e o vector de fecho) que foram calculadas usando aproximações iniciais aos parâmetros desconhecidos e valores observados, como imposição do processo de linearização. Dado que reduzimos um modelo não linear a um modelo linear, existe necessidade de realizar novas iterações, em particular se as aproximações iniciais aos parâmetros forem muito diferentes dos valores estimados. Em cada nova iteração, a linearização é feita em torno das estimativas dos parâmetros desconhecidos obtidos na iteração anterior e o processo iterativo deverá terminar quando as correcções às estimativas já obtidas forem nulas ou satisfizerem outro critério de precisão imposto. Dado que estamos apenas interessados nos valores finais, não existe necessidade de calcular as matrizes de variâncias-covariâncias em todas as iterações (minimizando desta forma o esforço computacional). Em síntese, para cada iteração *i*, o processo de ajustamento combinado é descrito pelas equações expressas na Tabela 3.

Tabela 3 – Equações para o ajustamento combinado.

$$\mathbf{A}_{\mathrm{m}\times\mathrm{u}} = \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{x}}\right|_{\boldsymbol{x}^{(\mathrm{i})},\boldsymbol{\ell}}$$

$$\mathbf{B}_{\mathrm{m}\times\mathrm{n}} = \left.\frac{\partial \mathbf{f}}{\partial \boldsymbol{\ell}}\right|_{\boldsymbol{x}^{(\mathrm{i})},\boldsymbol{\ell}}$$

$$\boldsymbol{w}_{\mathrm{m}\times 1}^{(\mathrm{i})} = \mathbf{f}\left(\boldsymbol{x}^{(\mathrm{i})},\boldsymbol{\ell},\mathbf{c}\right)$$

$$\hat{\delta}^{(\mathrm{i})} = -\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{w}$$

$$\hat{\boldsymbol{x}}^{(\mathrm{i})} = \boldsymbol{x}^{(\mathrm{i}-1)} + \hat{\delta}^{(\mathrm{i})}$$

$$\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}\hat{\delta} + \boldsymbol{w}\right)$$

$$\hat{\boldsymbol{v}} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}$$

$$\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \hat{\boldsymbol{v}}$$

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}$$

$$\mathbf{C}_{\hat{x}} = \mathbf{C}_{\hat{\delta}}$$

$$\mathbf{C}_{\hat{v}} = \sigma_0^2\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \mathbf{C}_{\ell} - \mathbf{C}_{\hat{v}}$$

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{\boldsymbol{v}}}{\mathrm{df}}$$

$$\hat{\mathbf{C}}_{\hat{\delta}} = \hat{\sigma}_0^2\left(\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}$$

$$\hat{\mathbf{C}}_{\hat{x}} = \hat{\mathbf{C}}_{\hat{\delta}}$$

$$\hat{\mathbf{C}}_{\hat{v}} = \hat{\sigma}_0^2\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)$$

$$\hat{\mathbf{C}}_{\hat{\ell}} = \mathbf{C}_{\ell} - \hat{\mathbf{C}}_{\hat{v}}$$

Ajustamento Paramétrico Não Linear

Consideremos de novo o caso do modelo matemático correspondente a um ajustamento paramétrico, mas vamos considerar que as relações entre os parâmetros desconhecidos e as observações são não lineares. O modelo matemático é dado por:

$$\mathbf{f}(\hat{x}) = \hat{\ell}\,, \tag{2.78}$$

onde **f** é um vector de funções não lineares.

Linearizando, temos:

$$\hat{\ell} \doteq \mathbf{f}(\mathbf{x}_0) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}}\right|_{x_0,\ell} (\hat{\mathbf{x}} - \mathbf{x}_0) \tag{2.79}$$

ou

$$\hat{\ell} \doteq \mathbf{f}(\mathbf{x}_0) + \mathbf{A}\delta\,. \tag{2.80}$$

Dado que

$$\hat{\ell} = \ell + v\,,$$

a Equação (2.80) pode ser escrita sob a forma:

$$v \doteq \mathbf{f}(\mathbf{x}_0) + \mathbf{A}\delta - \ell\,. \tag{2.81}$$

Fazendo

$$w = \mathbf{f}(\mathbf{x}_0) - \ell\,, \tag{2.82}$$

vem

$$v = w + \mathbf{A}\delta\,,$$

ou seja,

$$\boxed{\mathbf{A}\delta - v + w = \mathbf{0}}. \tag{2.83}$$

A equação anterior mostra que o ajustamento paramétrico para funções não lineares não é mais do que um caso particular do ajustamento combinado, com **B** = -**I**. A dedução das expressões para o vector de parâmetros desconhecidos e das observações ajustadas faz-se de forma idêntica à utilizada no ajustamento combinado.

Pretendemos minimizar a função

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} \,,$$

sujeita à condição expressa pela Equação (2.83). Introduzindo o vector de multiplicadores de Lagrange, a função a minimizar é escrita sob a forma

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{\boldsymbol{v}} + 2\hat{\mathbf{k}}^{\mathrm{T}} \left(\mathbf{A}\hat{\delta} - \hat{\boldsymbol{v}} + \boldsymbol{w} \right).$$

Calculando as derivadas parciais em relação a $\hat{\boldsymbol{v}}$ e a $\hat{\delta}$, obtemos respectivamente:

$$\frac{\partial \phi}{\partial \hat{\boldsymbol{v}}} = 2\hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_{\ell} - 2\hat{\mathbf{k}}^{\mathrm{T}} = \mathbf{0} \tag{2.84}$$

e

$$\frac{\partial \phi}{\partial \hat{\delta}} = 2\hat{\mathbf{k}}^{\mathrm{T}} \mathbf{A} = \mathbf{0}. \tag{2.85}$$

ou, de forma equivalente,

$$\mathbf{P}_{\ell} \hat{\boldsymbol{v}} - \hat{\mathbf{k}} = \mathbf{0} \tag{2.86}$$

$$\mathbf{A}^{\mathrm{T}} \hat{\mathbf{k}} = \mathbf{0}. \tag{2.87}$$

O sistema de equações a resolver é dado pelas equações (2.83), (2.86) e (2.87):

$$\begin{cases} \mathbf{A}\hat{\delta} - \hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0} \\ \mathbf{P}_{\ell} \hat{\boldsymbol{v}} - \hat{\mathbf{k}} = \mathbf{0} \\ \mathbf{A}^{\mathrm{T}} \hat{\mathbf{k}} = \mathbf{0} \end{cases}$$

Estas três equações podem ser combinadas na seguinte hipermatriz:

$$\left[\begin{array}{c:cc} \mathbf{P}_{\ell} & -\mathbf{I} & \mathbf{0} \\ \hdashline -\mathbf{I} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{0} \end{array} \right] \left[\begin{array}{c} \hat{\boldsymbol{v}} \\ \hdashline \hat{\mathbf{k}} \\ \hat{\delta} \end{array} \right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline \boldsymbol{w} \\ \mathbf{0} \end{array} \right] = \mathbf{0}. \tag{2.88}$$

Utilizando a Equação (2.57), obtemos:

$$\left[\begin{bmatrix}\mathbf{0} & \mathbf{A}\\ \mathbf{A}^{\mathrm{T}} & \mathbf{0}\end{bmatrix} - \begin{bmatrix}-\mathbf{I}\\ \mathbf{0}\end{bmatrix}\mathbf{P}_{\ell}^{-1}\begin{bmatrix}-\mathbf{I} & \mathbf{0}\end{bmatrix}\right]\begin{bmatrix}\hat{\mathbf{k}}\\ \hat{\delta}\end{bmatrix} + \left[\begin{bmatrix}\boldsymbol{w}\\ \mathbf{0}\end{bmatrix} - \begin{bmatrix}-\mathbf{I}\\ \mathbf{0}\end{bmatrix}\mathbf{P}_{\ell}^{-1}\begin{bmatrix}\mathbf{0}\end{bmatrix}\right] = \mathbf{0}\,.$$

$$\begin{bmatrix}-\mathbf{P}_{\ell}^{-1} & \mathbf{A}\\ \mathbf{A}^{\mathrm{T}} & \mathbf{0}\end{bmatrix}\begin{bmatrix}\hat{\mathbf{k}}\\ \hat{\delta}\end{bmatrix} + \begin{bmatrix}\boldsymbol{w}\\ \mathbf{0}\end{bmatrix} = \mathbf{0}\,. \tag{2.89}$$

Para determinar $\hat{\delta}$ particiona-se a hipermatriz de modo a que $\mathbf{X}_1 = \hat{\mathbf{k}}$:

$$\left[\begin{array}{c:c}-\mathbf{P}_{\ell}^{-1} & \mathbf{A}\\ \hdashline \mathbf{A}^{\mathrm{T}} & \mathbf{0}\end{array}\right]\left[\begin{array}{c}\hat{\mathbf{k}}\\ \hdashline \hat{\delta}\end{array}\right] + \left[\begin{array}{c}\boldsymbol{w}\\ \hdashline \mathbf{0}\end{array}\right] = \mathbf{0}\,. \tag{2.90}$$

Aplicando de novo a Equação (2.57), tem-se:

$$\boxed{\hat{\delta} = -\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\,\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\,\boldsymbol{w}}\,. \tag{2.91}$$

Para obtermos $\hat{\mathbf{k}}$ utilizamos a primeira equação de (2.90):

$$\boxed{\hat{\mathbf{k}} = \mathbf{P}_{\ell}\left(\mathbf{A}\hat{\delta} + \boldsymbol{w}\right)}\,. \tag{2.92}$$

O vector de resíduos $\hat{v}$ obtém-se da primeira equação de (2.88):

$$\hat{v} = \mathbf{P}_{\ell}^{-1}\hat{\mathbf{k}}\,, \tag{2.93}$$

ou, de forma equivalente:

$$\boxed{\hat{v} = \mathbf{A}\hat{\delta} + \boldsymbol{w}}\,. \tag{2.94}$$

Repare-se na semelhança existente entre estas equações e as estabelecidas para o caso linear. A diferença mais significativa é que, contrariamente ao caso das funções lineares, o que se obtém agora são correcções aos parâmetros iniciais e não os parâmetros propriamente ditos. Além disso, há que ter em consideração a necessidade de eventualmente ser necessário efectuar mais do que uma iteração. As estimativas para o vector de parâmetros e para o vector de observações ajustadas são:

$$\boxed{\hat{x} = x_0 + \hat{\delta}}$$

$$\boxed{\hat{\ell} = \ell + \hat{v}}\,.$$

O vector de observações ajustadas pode ser escrito de forma mais explícita, útil para derivar a respectiva matriz de variâncias-covariâncias:

$$\boxed{\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\boldsymbol{w} + \boldsymbol{w}}. \tag{2.95}$$

Matrizes de variâncias-covariâncias

As matrizes de variâncias-covariâncias são obtidas usando um procedimento análogo ao utilizado no ajustamento combinado. Temos

Para obter a **matriz de variâncias-covariâncias para** $\hat{\delta}$, tem-se:

$$\mathbf{C}_{\hat{\delta}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}}, \tag{2.96}$$

sendo

$$\mathbf{J} = \frac{\partial\hat{\delta}}{\partial\boldsymbol{\ell}} = \left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell},$$

dado que $\boldsymbol{w} = \mathbf{f}(x_0) - \boldsymbol{\ell}$. Como $\mathbf{C}_{\boldsymbol{\ell}} = \sigma_0^2\mathbf{P}_{\boldsymbol{\ell}}^{-1}$, obtemos, após substituição:

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\cancel{\mathbf{P}_{\ell}^{-1}}\cancel{\mathbf{P}_{\ell}}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\right)$$

$$\boxed{\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}}. \tag{2.97}$$

Para a matriz de variâncias-covariâncias das observações ajustadas, temos:

$$\mathbf{C}_{\hat{\ell}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}}$$

Usando a equação (2.95) obtemos:

$$\mathbf{J} = \frac{\partial\hat{\boldsymbol{\ell}}}{\partial\boldsymbol{\ell}} = \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell},$$

logo

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\cancel{\mathbf{P}_{\ell}^{-1}}\cancel{\mathbf{P}_{\ell}}\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right). \tag{2.98}$$

De forma idêntica se demonstra que

$$\boxed{\mathbf{C}_{\hat{v}} = \sigma_0^2\left(\mathbf{P}_\ell^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right)}, \tag{2.99}$$

pelo que, tal como anteriormente, se tem a relação

$$\boxed{\mathbf{C}_{\hat{\ell}} = \mathbf{C}_\ell - \mathbf{C}_{\hat{v}}}. \tag{2.100}$$

As equações para o ajustamento paramétrico não linear são apresentadas na Tabela 4.

Tabela 4 – Equações para o ajustamento paramétrico não linear.

$$\mathbf{A}_{\mathrm{m}\times\mathrm{u}} = \left.\frac{\partial \mathbf{f}}{\partial \hat{\boldsymbol{x}}}\right|_{x^{(\mathrm{i})},\ell}$$

$$\boldsymbol{w}_{\mathrm{m}\times 1}^{(\mathrm{i})} = \mathbf{f}\left(\boldsymbol{x}^{(\mathrm{i})}\right) - \boldsymbol{\ell}$$

$$\hat{\delta}^{(\mathrm{i})} = -\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\boldsymbol{w}$$

$$\hat{\boldsymbol{x}}^{(\mathrm{i})} = \boldsymbol{x}^{(\mathrm{i}-1)} + \hat{\delta}^{(\mathrm{i})}$$

$$\hat{\boldsymbol{v}} = \mathbf{A}\hat{\delta}^{(\mathrm{i})} + \boldsymbol{w}$$

$$\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} + \hat{\boldsymbol{v}}$$

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}$$

$$\mathbf{C}_{\hat{x}} = \mathbf{C}_{\hat{\delta}}$$

$$\mathbf{C}_{\hat{v}} = \sigma_0^2\left(\mathbf{P}_\ell^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \mathbf{C}_\ell - \mathbf{C}_{\hat{v}}$$

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_\ell\hat{\boldsymbol{v}}}{\mathrm{df}}$$

$$\hat{\mathbf{C}}_{\hat{\delta}} = \hat{\sigma}_0^2\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}$$

$$\hat{\mathbf{C}}_{\hat{v}} = \hat{\sigma}_0^2\left(\mathbf{P}_\ell^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_\ell\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right)$$

$$\hat{\mathbf{C}}_{\hat{\ell}} = \mathbf{C}_\ell - \hat{\mathbf{C}}_{\hat{v}}$$

Ajustamento Condicional Não Linear

Para o ajustamento condicional, o modelo matemático é do tipo:

$$\mathbf{f}\left(\hat{\boldsymbol{\ell}}\right)=\mathbf{0}\,. \tag{2.101}$$

Para modelos não lineares, a linearização conduz a

$$\mathbf{f}\left(\hat{\boldsymbol{\ell}}\right)\doteq\mathbf{f}\left(\boldsymbol{\ell}\right)+\left.\frac{\partial\mathbf{f}}{\partial\hat{\boldsymbol{\ell}}}\right|_{\boldsymbol{\ell}}\left(\hat{\boldsymbol{\ell}}-\boldsymbol{\ell}\right)=\mathbf{0}\,, \tag{2.102}$$

ou, em forma matricial,

$$\boxed{\mathbf{B}\boldsymbol{v}+\boldsymbol{w}=\mathbf{0}}, \tag{2.103}$$

onde

$$\mathbf{B}_{\mathrm{m\times n}}=\left.\frac{\partial\mathbf{f}}{\partial\hat{\boldsymbol{\ell}}}\right|_{\boldsymbol{\ell}}$$

$$\boldsymbol{v}=\hat{\boldsymbol{\ell}}-\boldsymbol{\ell}$$

$$\boldsymbol{w}=\mathbf{f}\left(\boldsymbol{\ell}\right).$$

Analisando a Equação (2.103), reconhece-se que o ajustamento condicional não linear é um caso particular do ajustamento combinado, onde **A=0**.

Pretendemos minimizar a função

$$\phi=\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{\boldsymbol{v}}\,.$$

Introduzindo o vector de multiplicadores de Lagrange, a função a minimizar será

$$\phi=\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{\boldsymbol{v}}+2\hat{\mathbf{k}}^{\mathrm{T}}\left(\mathbf{B}\hat{\boldsymbol{v}}+\boldsymbol{w}\right).$$

Tem-se

$$\frac{\partial\phi}{\partial\hat{\boldsymbol{v}}}=2\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}+2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{B}=\mathbf{0}\,,$$

ou seja,

$$\mathbf{P}_{\ell}\hat{\boldsymbol{v}}+\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}=\mathbf{0}\,.$$

Pretendemos resolver o sistema

$$\begin{bmatrix} \mathbf{P}_\ell & \mathbf{B}^\mathrm{T} \\ \mathbf{B} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{v}} \\ \hat{\mathbf{k}} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{w} \end{bmatrix} = \mathbf{0} \ . \tag{2.104}$$

Aplicando a Equação (2.57), obtém-se

$$-\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\right)\hat{\mathbf{k}} + \boldsymbol{w} = \mathbf{0} \tag{2.105}$$

$$\boxed{\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\right)^{-1}\boldsymbol{w}}. \tag{2.106}$$

O vector de resíduos obtém-se a partir da primeira Equação do sistema (2.104):

$$\mathbf{P}_\ell\,\hat{\boldsymbol{v}} + \mathbf{B}^\mathrm{T}\hat{\mathbf{k}} = \mathbf{0}$$

$$\boxed{\hat{\boldsymbol{v}} = -\mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\hat{\mathbf{k}}}, \tag{2.107}$$

pelo que o vector de observações ajustadas é

$$\boxed{\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} - \mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\mathbf{M}^{-1}\boldsymbol{w}}. \tag{2.108}$$

Matrizes de variâncias-covariâncias

Para a **matriz de variâncias-covariâncias de $\hat{\boldsymbol{\ell}}$**, temos:

$$\mathbf{C}_{\hat{\ell}} = \mathbf{J}\mathbf{C}_\ell\mathbf{J}^\mathrm{T}$$

Usando a equação (2.108) obtemos

$$\mathbf{J} = \frac{\partial\hat{\boldsymbol{\ell}}}{\partial\boldsymbol{\ell}} = \mathbf{I} - \mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\mathbf{M}^{-1}\mathbf{B}\,,$$

pelo que, utilizando igualmente a relação $\mathbf{C}_\ell = \sigma_0^2\mathbf{P}_\ell^{-1}$, se tem

$$\mathbf{C}_{\hat{\ell}} = \sigma_0^2\left(\mathbf{I} - \mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\mathbf{M}^{-1}\mathbf{B}\right)\left(\mathbf{P}_\ell^{-1} - \mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}\right)$$

ou seja, após manipulação e simplificação,

$$\boxed{\mathbf{C}_{\hat{\ell}} = \sigma_0^2\mathbf{P}_\ell^{-1} - \sigma_0^2\mathbf{P}_\ell^{-1}\mathbf{B}^\mathrm{T}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_\ell^{-1}} \tag{2.109}$$

A **matriz de variâncias-covariâncias para $\hat{\boldsymbol{v}}$**, tem-se

$$\mathbf{C}_{\hat{v}} = \mathbf{J}\mathbf{C}_\ell\mathbf{J}^\mathrm{T}$$

com

$$\mathbf{J} = \frac{\partial \hat{\boldsymbol{v}}}{\partial \boldsymbol{\ell}} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}.$$

Substituindo e simplificando, obtemos

$$\boxed{\mathbf{C}_{\hat{v}} = \sigma_0^2\left(\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)}. \tag{2.110}$$

Das Equações (2.109) e (2.110) conclui-se que

$$\boxed{\mathbf{C}_{\hat{\ell}} = \mathbf{C}_{\ell} - \mathbf{C}_{\hat{v}}}. \tag{2.77}$$

As equações para o ajustamento condicional não linear são as constantes na Tabela 5.

Tabela 5 – Equações para o ajustamento condicional não linear.

$$\mathbf{B}_{m\times n} = \left.\frac{\partial \mathbf{f}}{\partial \hat{\boldsymbol{\ell}}}\right|_{\hat{\boldsymbol{\ell}}=\boldsymbol{\ell}^{(i)}}$$

$$\boldsymbol{w}_{m\times 1}^{(i)} = \mathbf{f}\left(\boldsymbol{\ell}^{(i)}\right)$$

$$\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{w}$$

$$\hat{\boldsymbol{v}}^{(i)} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}$$

$$\hat{\boldsymbol{\ell}}^{(i)} = \boldsymbol{\ell}^{(i-1)} + \hat{\boldsymbol{v}}^{(i)}$$

$$\mathbf{C}_{\hat{v}} = \sigma_0^2\left(\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{B}\mathbf{P}_{\ell}^{-1}\right)$$

$$\mathbf{C}_{\hat{\ell}} = \mathbf{C}_{\ell} - \mathbf{C}_{\hat{v}}$$

$$\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{\boldsymbol{v}}}{\mathrm{df}}$$

$$\hat{\mathbf{C}}_{\hat{v}} = \hat{\sigma}_0^2\left(\mathbf{P}_{\ell}^{-1} - \mathbf{A}\left(\mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\right)$$

$$\hat{\mathbf{C}}_{\hat{\ell}} = \mathbf{C}_{\ell} - \hat{\mathbf{C}}_{\hat{v}}$$

Notas Finais

Ao longo deste capítulo provou-se que a escolha da variância *a priori* não influencia quer a estimativa obtida para os parâmetros desconhecidos, $\hat{\boldsymbol{x}}$, quer a estimativa da matriz de variâncias-covariâncias deste vector, $\hat{\mathbf{C}}_{\hat{x}}$. Assim, estas quantidades podem ser obtidas assumindo qualquer peso relativo entre as observações. No entanto, a escolha do valor da variância *a priori* influencia a matriz das equações normais – como é o caso, por exemplo, da matriz de equações normais para o ajustamento paramétrico linear, $\mathbf{N} = \mathbf{A}^{\mathrm{T}}\mathbf{P}_{\ell}\,\mathbf{A}$, e a variância *a posteriori*, $\hat{\sigma}_0^2 = \dfrac{\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{\boldsymbol{v}}}{\mathrm{df}}$. Assim, que papel desempenha o factor de variância *a priori* no processo de ajustamento? Existem várias situações particulares em que a escolha do valor do σ_0^2 se revela de grande interesse, tais como:

- evitar instabilidades numéricas no processo de inversão da matriz das equações normais;

  Uma escolha eficiente de σ_0^2 permite modificar a magnitude dos elementos das equações normais e fazer com que os elementos de **N** sejam, em média, tão próximos da unidade quanto possível, ou simplesmente mais convenientes para que o processo numérico da sua inversão seja mais preciso.

- testar os valores assumidos para a matriz de variâncias-covariâncias das observações;
- testar a consistência do modelo matemático com as observações.

Vamos supor que na realização de um processo de ajustamento, e na falta de um melhor critério, se assumiu que $\sigma_0^2 = 1$. Se a matriz de covariâncias das observações for adequada e o modelo matemático representar de forma correcta o problema, a razão entre a variância *a priori* e a variância *a posteriori*, obtida após o ajustamento, deverá ser próxima da unidade. O que pode conduzir a valores anormais para a razão entre variâncias? Como resolver essas situações?

Existem 3 situações que conduzem à discrepância entre os valores das variâncias *a priori* e *a posteriori*: (1) formulação pessimista ou optimista da matriz de variâncias-covariâncias das observações; (2) erros sistemáticos nas observações; (3) inconsistência entre modelo matemático e a realidade fisico-geométrica.

Na impossibilidade ou dificuldade de encontrar uma matriz de variâncias-covariâncias das observações que seja adequada, o primeiro caso pode ser solucionado usando uma variância *a priori* que corrija essa situação, com base na informação dada pela variância *a posteriori*. Na generalidade dos casos, o problema é solucionado após esta reformulação do sistema de pesos.

Se a variância *a posteriori* obtida é estatisticamente diferente da variância *a priori,* após uma utilização adequada da matriz de pesos (considerando experiências prévias, especificações de fabrico para determinado instrumento, etc.), há que examinar o modelo matemático e verificar se ele representa efectivamente o problema (terceiro caso).

Se existirem erros sistemáticos (ou enganos) nas observações, estes erros manifestam-se como desvios ao modelo matemático. Na tentativa de ajustar todas as observações existentes, os resíduos tomarão valores anormais, anormalidade que se reflectirá no cálculo da variância *a posteriori* (segundo caso). De facto, a análise de resíduos constitui uma ferramenta fundamental na detecção de erros sistemáticos e no processo de decisão da qualidade da solução obtida.

Qualquer dos métodos de ajustamento referidos conduz-nos a um conjunto de equações normais. O que existe afinal de "normal" num sistema de equações normais? A resposta a esta questão é a essência do ajustamento por mínimos quadrados. A aplicação do critério dos mínimos quadrados transformou um sistema indeterminado ("anormal") num sistema consistente, para o qual existe uma única solução ("normal").

Finalmente um comentário sobre a escolha do método a utilizar. A escolha da técnica de ajustamento é fundamentalmente condicionada pelo modelo matemático e pelo tipo de aplicação. O método paramétrico é vantajoso em implementações de cálculo automático, uma vez que cada observação gera uma única equação. Para além disso, conduz-nos directamente aos valores dos parâmetros desconhecidos. Há umas décadas, o ajustamento condicional era preferido, uma vez que necessitava de um menor número de equações. Com a introdução dos computadores, não existem razões fortes para a sua escolha. Dado que não conduz directamente aos parâmetros desconhecidos, este método revela-se algo "incompleto", requerendo um cálculo adicional. É particularmente útil no ajustamento de pequenas redes de nivelamento (gravimetria) e em situações em que se está mais interessado na obtenção de observações ajustadas, mas pouco usado actualmente. As situações que requerem a utilização do método combinado deixam-nos, à partida, sem hipótese de escolha. Prova-se, no entanto, que é possível reformular um problema de tal forma que o ajustamento paramétrico possa ser aplicado.

3. MÉTODOS AVANÇADOS DE AJUSTAMENTO

Embora os métodos de ajustamento "clássicos" resolvam uma grande parte dos problemas de ajustamento em Ciências Geográficas, há por vezes necessidade de alargar os modelos matemáticos, através de partição dos modelos básicos. Exemplos de situações deste tipo incluem: (1) casos em que o modelo matemático contém parâmetros que têm que ser estimados, mas para os quais não existe interesse particular nos valores que para eles venham a ser obtidos; (2) casos em que pretendemos incluir observações adicionais, obtidas por instrumentação diferente ou noutra época de observações, por exemplo; (3) casos em que se tem informação funcional adicional; (4) casos em que existe algum conhecimento sobre os valores dos parâmetros a determinar.

Nesta secção, parte-se do pressuposto que os modelos matemáticos correspondem sempre ao caso mais geral, ou seja, são do tipo implícito. As expressões obtidas podem depois ser facilmente simplificadas para modelos do tipo explícito (paramétrico), como acontece na adaptação do ajustamento combinado a um ajustamento paramétrico não linear.

Ajustamento Com Parâmetros Inúteis

Consideremos o caso em que o modelo matemático inclui simultaneamente parâmetros desconhecidos de interesse e outros parâmetros, que apenas têm que ser determinados por imposição do modelo matemático – parâmetros inúteis (ou parâmetros adicionais). Exemplos de parâmetros inúteis: os rumos do zero do limbo, a incluir nas equações de observação de direcções azimutais; as ambiguidades de ciclo, a incluir nas equações de observação de fase, em sistemas de posicionamento por satélite.

Designemos por $\boldsymbol{x}_1$ o vector de parâmetros inúteis, de dimensões $u_1 \times 1$, e por $\boldsymbol{x}_2$ o vector dos parâmetros de interesse, de dimensões $u_2 \times 1$. O modelo matemático, que assumimos como não linear e do tipo combinado, é dado por

$$\mathbf{f}\left(\hat{x}_1, \hat{x}_2, \hat{\ell}\right) = \mathbf{0}, \qquad (3.1)$$

ou seja, um conjunto de m equações de observação, envolvendo n observações e u parâmetros, sendo $u = u_1 + u_2$:

$$m = u + df.$$

Linearizando em torno dos valores aproximados x_1^0, x_2^0, e ℓ, obtemos

$$\mathbf{f}\left(\hat{x}_1, \hat{x}_2, \hat{\ell}\right) \doteq \mathbf{f}\left(x_1^0, x_2^0, \ell\right) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}_1}\right|_{x_1^0, x_2^0, \ell} \left(\hat{x}_1 - x_1^0\right) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}_2}\right|_{x_1^0, x_2^0, \ell} \left(\hat{x}_2 - x_2^0\right) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\ell}}\right|_{x_1^0, x_2^0, \ell} \left(\hat{\ell} - \ell\right) = \mathbf{0}$$

ou, em forma matricial,

$$\boxed{\mathbf{A}_1 \hat{\delta}_1 + \mathbf{A}_2 \hat{\delta}_2 + \mathbf{B}\hat{v} + w = \mathbf{0}}, \qquad (3.2)$$

onde

$${}_{m\times u_1}\mathbf{A}_1 = \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}_1}\right|_{x_1^0, x_2^0, \ell}$$ primeira matriz de configuração, relativa aos parâmetros inúteis

$${}_{m\times u_2}\mathbf{A}_2 = \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}_2}\right|_{x_1^0, x_2^0, \ell}$$ primeira matriz de configuração, relativa aos parâmetros de interesse

$${}_{m\times n}\mathbf{B} = \left.\frac{\partial \mathbf{f}}{\partial \hat{\ell}}\right|_{x_1^0, x_2^0, \ell}$$ segunda matriz de configuração

$${}_{u_1\times 1}\hat{\delta}_1 = \left(\hat{x}_1 - x_1^0\right)$$ vector de correcções às estimativas iniciais dos parâmetros inúteis

$${}_{u_2\times 1}\hat{\delta}_2 = \left(\hat{x}_2 - x_2^0\right)$$ vector de correcções às estimativas iniciais dos parâmetros de interesse

$${}_{n\times 1}\hat{v} = \hat{\ell} - \ell$$ vector de resíduos

$${}_{m\times 1}w = \mathbf{f}\left(x_1^0, x_2^0, \ell\right)$$ vector de fecho

De notar que este modelo matemático é equivalente à partição da matriz **A** e do vector x (δ), correspondente à versão linearizada do modelo matemático do ajustamento combinado (note-se a conformidade para a multiplicação das matrizes particionadas):

$$\boxed{\mathbf{A}\hat{\delta}+\mathbf{B}\hat{v}+w=\mathbf{0}},$$

onde

$$\mathbf{A}=\left[\mathbf{A}_1 \;\vdots\; \mathbf{A}_2\right]$$

$$\hat{\delta}=\begin{bmatrix}\hat{\delta}_1\\ \hline \hat{\delta}_2\end{bmatrix}.$$

Pretende-se minimizar a função

$$\phi=\hat{v}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{v}$$

com o constrangimento

$$\mathbf{A}_1\hat{\delta}_1+\mathbf{A}_2\hat{\delta}_2+\mathbf{B}\hat{v}+w=\mathbf{0}.$$

Introduzindo o vector de multiplicadores de Lagrange, a função de variação ϕ que pretendemos minimizar é

$$\phi=\hat{v}^{\mathrm{T}}\mathbf{P}_{\ell}\hat{v}+2\hat{\mathbf{k}}^{\mathrm{T}}\left(\mathbf{A}_1\hat{\delta}_1+\mathbf{A}_2\hat{\delta}_2+\mathbf{B}\hat{v}+w\right). \tag{3.3}$$

Para minimizar esta função, começamos por igualar a zero as diversas derivadas parciais. Derivando em ordem a $\hat{v}$, obtemos

$$\frac{\partial\phi}{\partial\hat{v}}=2\hat{v}^{\mathrm{T}}\mathbf{P}_{\ell}+2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{B}=\mathbf{0},$$

ou, transpondo ambos os membros da equação e simplificando,

$$\boxed{\mathbf{P}_{\ell}\hat{v}+\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}=\mathbf{0}}. \tag{3.4}$$

De forma idêntica, temos

$$\frac{\partial\phi}{\partial\hat{\delta}_1}=2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{A}_1=\mathbf{0},$$

ou

$$\boxed{\mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}}. \tag{3.5}$$

Por fim, obtemos

$$\frac{\partial \phi}{\partial \hat{\delta}_2} = 2\hat{\mathbf{k}}^{\mathrm{T}}\mathbf{A}_2 = \mathbf{0},$$

ou

$$\boxed{\mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}}. \tag{3.6}$$

Pretendemos agora encontrar a solução para o sistema constituído pelas Equações (3.2), (3.4), (3.5) e (3.6):

$$\begin{cases} \mathbf{A}_1\hat{\delta}_1 + \mathbf{A}_2\hat{\delta}_2 + \mathbf{B}\hat{v} + \boldsymbol{w} = \mathbf{0} \\ \mathbf{P}_\ell\hat{v} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \\ \mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \\ \mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \end{cases} \tag{3.7}$$

que pode ser expresso na forma matricial

$$\begin{bmatrix} \mathbf{P}_\ell & \mathbf{B}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \mathbf{B} & \mathbf{0} & \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{A}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \hat{v} \\ \hat{\mathbf{k}} \\ \hat{\delta}_1 \\ \hat{\delta}_2 \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \boldsymbol{w} \\ \mathbf{0} \\ \mathbf{0} \end{bmatrix} = \mathbf{0}. \tag{3.8}$$

Tal como no ajustamento combinado, a solução deste sistema de equações é obtida utilizando as propriedades da partição por blocos de matrizes. Assim, estabelecendo a seguinte partição por blocos:

$$\left[\begin{array}{c:ccc} \mathbf{P}_\ell & \mathbf{B}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \hdashline \mathbf{B} & \mathbf{0} & \mathbf{A}_1 & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{A}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v} \\ \hdashline \hat{\mathbf{k}} \\ \hat{\delta}_1 \\ \hat{\delta}_2 \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline \boldsymbol{w} \\ \mathbf{0} \\ \mathbf{0} \end{array}\right] = \mathbf{0}$$

e utilizando a Equação (2.57), podemos eliminar $\hat{v}$. O sistema de equações resultante é dado por

$$\left[\begin{array}{c:cc} -\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}} & \mathbf{A}_1 & \mathbf{A}_2 \\ \hdashline \mathbf{A}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \end{array}\right]\left[\begin{array}{c} \hat{\mathbf{k}} \\ \hdashline \hat{\delta}_1 \\ \hat{\delta}_2 \end{array}\right]+\left[\begin{array}{c} \boldsymbol{w} \\ \hdashline \mathbf{0} \\ \mathbf{0} \end{array}\right]=\mathbf{0}. \tag{3.9}$$

A partição anterior permite eliminar $\hat{\mathbf{k}}$, obtendo-se

$$\left[\begin{array}{c:c} \mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1 & \mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 \\ \hdashline \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1 & \mathbf{A}_2^{\mathrm{T}-1}\mathbf{A}_2 \end{array}\right]\left[\begin{array}{c} \hat{\delta}_1 \\ \hdashline \hat{\delta}_2 \end{array}\right]+\left[\begin{array}{c} \mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w} \\ \hdashline \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w} \end{array}\right]=\mathbf{0}, \tag{3.10}$$

onde $\mathbf{M}=\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}$. Em alternativa, podemos adoptar a representação

$$\left[\begin{array}{c:c} \mathbf{N}_{11} & \mathbf{N}_{12} \\ \hdashline \mathbf{N}_{21} & \mathbf{N}_{22} \end{array}\right]\left[\begin{array}{c} \hat{\delta}_1 \\ \hdashline \hat{\delta}_2 \end{array}\right]+\left[\begin{array}{c} \mathbf{u}_1 \\ \hdashline \mathbf{u}_2 \end{array}\right]=\mathbf{0}, \tag{3.11}$$

onde

$$\mathbf{N}_{11}=\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1$$

$$\mathbf{N}_{12}=\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2$$

$$\mathbf{N}_{21}=\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1$$

$$\mathbf{N}_{22}=\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2$$

$$\mathbf{u}_1=\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}$$

$$\mathbf{u}_2=\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}$$

Eliminando $\hat{\delta}_1$, obtemos:

$$\boxed{\hat{\delta}_2=-\left(\mathbf{N}_{22}-\mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)^{-1}\left(\mathbf{u}_2-\mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{u}_1\right)}. \tag{3.12}$$

Como se pode constatar, chegámos a uma expressão que nos dá as correcções às estimativas dos parâmetros desconhecidos de interesse, sem que, para tal, intervenha qualquer conhecimento dos parâmetros inúteis; no entanto, o efeito dos parâmetros eliminados é rigorosamente mantido na restante solução e podem ser obtidos utilizando a primeira equação do sistema de equações (3.11):

$$\mathbf{N}_{11}\hat{\delta}_1 + \mathbf{N}_{12}\hat{\delta}_2 + \mathbf{u}_1 = \mathbf{0}\,.$$

Desta equação, obtemos

$$\boxed{\hat{\delta}_1 = -\mathbf{N}_{11}^{-1}\left(\mathbf{N}_{12}\hat{\delta}_2 + \mathbf{u}_1\right)}\,. \tag{3.13}$$

O vector de multiplicadores de Lagrange obtém-se da primeira equação do sistema (3.9):

$$-\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} + \mathbf{A}_1\hat{\delta}_1 + \mathbf{A}_2\hat{\delta}_2 + \boldsymbol{w} = \mathbf{0}\,,$$

donde

$$\boxed{\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}_1\hat{\delta}_1 + \mathbf{A}_2\hat{\delta}_2 + \boldsymbol{w}\right)}. \tag{3.14}$$

O vector dos resíduos obtém-se da primeira equação do sistema (3.8):

$$\mathbf{P}_{\ell}\hat{\boldsymbol{v}} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}\,,$$

pelo que

$$\boxed{\hat{\boldsymbol{v}} = -\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}}. \tag{3.15}$$

Uma vez obtidos $\hat{\delta}_1$, $\hat{\delta}_2$, e $\hat{\boldsymbol{v}}$, tem-se:

$$\boxed{\hat{\boldsymbol{x}}_1 = \boldsymbol{x}_1^0 + \hat{\delta}_1} \tag{3.16}$$

$$\boxed{\hat{\boldsymbol{x}}_2 = \boldsymbol{x}_2^0 + \hat{\delta}_2} \tag{3.17}$$

$$\boxed{\hat{\boldsymbol{\ell}} = \boldsymbol{\ell} - \mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}}. \tag{3.18}$$

Matriz de variâncias-covariâncias dos parâmetros de interesse

A dedução das expressões das matrizes de variâncias-covariâncias das diversas variáveis faz-se novamente recorrendo à lei de propagação de variâncias-covariâncias. Neste texto, vamos simplesmente apresentar a matriz de variâncias-covariâncias para $\hat{\delta}_2$.

A expressão para $\hat{\delta}_2$ pode ser escrita sob a forma

$$\hat{\boldsymbol{\delta}}_2 = -\left(\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)^{-1}\left(\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\right)\boldsymbol{w}\,.$$

Tem-se

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \mathbf{J}\mathbf{C}_{\boldsymbol{\ell}}\mathbf{J}^{\mathrm{T}}\,,$$

onde

$$\mathrm{J} = \frac{\partial\hat{\boldsymbol{\delta}}_2}{\partial\boldsymbol{\ell}} = -\left(\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)^{-1}\left(\mathbf{A}_2^{\mathrm{T}}\mathrm{M}^{-1} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathrm{M}^{-1}\right)\mathbf{B}$$

Introduzindo uma matriz auxiliar $\mathbf{Z} = \left(\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)$, tem-se

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_2} = \sigma_0^2\mathbf{Z}^{-1}\left(\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\boldsymbol{\ell}}^{-1} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\mathbf{P}_{\boldsymbol{\ell}}^{-1}\right)\left(\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{N}_{21}^{\mathrm{T}}\right)\mathbf{Z}^{-1}\,.$$

Desenvolvendo e simplificando, vem:

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_2} = \sigma_0^2\mathbf{Z}^{-1}\left(\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{N}_{21}^{\mathrm{T}} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\underbrace{\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2}_{\mathbf{N}_{21}^{\mathrm{T}}} + \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\underbrace{\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1}_{\mathbf{N}_{11}}\mathbf{N}_{11}^{-1}\mathbf{N}_{21}^{\mathrm{T}}\right)\mathbf{Z}^{-1}$$

Mas

$$\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{N}_{21}^{\mathrm{T}} = \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{N}_{12} = \mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12} = \mathbf{Z}$$

logo

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_2} = \sigma_0^2\mathbf{Z}^{-1} = \sigma_0^2\left(\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)^{-1}$$

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_2} = \sigma_0^2\mathbf{Z}^{-1} = \sigma_0^2\left(\mathbf{N}_{22} - \mathbf{N}_{21}\mathbf{N}_{11}^{-1}\mathbf{N}_{12}\right)^{-1}$$

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}_2} = \sigma_0^2\left(\mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2 - \mathbf{A}_2^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_2\right)^{-1}} \qquad (3.19)$$

A variância *a posteriori* é dada por

$$\boxed{\hat{\sigma}_0^2 = \frac{\hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\hat{\boldsymbol{v}}}{\mathrm{df}}}\,, \qquad (3.20)$$

onde df=n-(u_1+u_2).

Ajustamento com Observações Adicionais

Consideremos o caso em que existem dois conjuntos de observações, que podem ter sido obtidas em épocas diferentes ou com instrumentação diferente, e que se pretende obter estimativas para os parâmetros desconhecidos com base nos dois conjuntos de equações. O modelo matemático é do tipo:

$$\begin{cases} \mathbf{f}_1\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_1\right)=\mathbf{0} \\ \mathbf{f}_2\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_2\right)=\mathbf{0} \end{cases}$$

onde $\boldsymbol{\ell}_1$ designa o vector de observações iniciais, de dimensões $\mathrm{n}_1 \times 1$, e $\boldsymbol{\ell}_2$ o vector de observações adicionais, de dimensões $\mathrm{n}_2 \times 1$, e aos quais associamos as matrizes de variâncias-covariâncias $\mathbf{C}_{\boldsymbol{\ell}_1}$ e $\mathbf{C}_{\boldsymbol{\ell}_2}$.

Linearizando as duas funções, usando séries de Taylor, em torno de valores aproximados $\boldsymbol{x}^0$, e dos valores observados $\boldsymbol{\ell}_1$ e $\boldsymbol{\ell}_2$, obtemos:

$$\mathbf{f}_1\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_1\right)\doteq\mathbf{f}_1\left(\boldsymbol{x}^0,\boldsymbol{\ell}_1\right)+\left.\frac{\partial\mathbf{f}_1}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_1}\left(\hat{\boldsymbol{x}}-\boldsymbol{x}^0\right)+\left.\frac{\partial\mathbf{f}_1}{\partial\hat{\boldsymbol{\ell}}_1}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_1}\left(\hat{\boldsymbol{\ell}}_1-\boldsymbol{\ell}_1\right)=\mathbf{0}$$

$$\mathbf{f}_2\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_2\right)\doteq\mathbf{f}_2\left(\boldsymbol{x}^0,\boldsymbol{\ell}_2\right)+\left.\frac{\partial\mathbf{f}_2}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_2}\left(\hat{\boldsymbol{x}}-\boldsymbol{x}^0\right)+\left.\frac{\partial\mathbf{f}_2}{\partial\hat{\boldsymbol{\ell}}_2}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_2}\left(\hat{\boldsymbol{\ell}}_2-\boldsymbol{\ell}_2\right)=\mathbf{0}$$

ou, em forma matricial:

$$\boxed{\mathbf{A}_1\hat{\boldsymbol{\delta}}+\mathbf{B}_1\hat{\boldsymbol{v}}_1+\boldsymbol{w}_1=\mathbf{0}} \tag{3.21}$$

$$\boxed{\mathbf{A}_2\hat{\boldsymbol{\delta}}+\mathbf{B}_2\hat{\boldsymbol{v}}_2+\boldsymbol{w}_2=\mathbf{0}}, \tag{3.22}$$

onde

$\underset{\mathrm{m}_1\times\mathrm{u}}{\mathbf{A}_1}=\left.\dfrac{\partial\mathbf{f}_1}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_1}$ primeira matriz de configuração, relativa ao primeiro conjunto de observações

$\underset{\mathrm{m}_2\times\mathrm{u}}{\mathbf{A}_2}=\left.\dfrac{\partial\mathbf{f}_2}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_2}$ primeira matriz de configuração, relativa ao segundo conjunto de observações

$_{m_1 \times n_1}\mathbf{B}_1 = \left.\dfrac{\partial \mathbf{f}_1}{\partial \hat{\boldsymbol{\ell}}_1}\right|_{\boldsymbol{x}^0, \boldsymbol{\ell}_1}$ segunda matriz de configuração, relativa ao primeiro conjunto de observações

$_{m_2 \times n_2}\mathbf{B}_2 = \left.\dfrac{\partial \mathbf{f}_2}{\partial \hat{\boldsymbol{\ell}}_2}\right|_{\boldsymbol{x}^0, \boldsymbol{\ell}_2}$ segunda matriz de configuração, relativa ao segundo conjunto de observações

$_{u \times 1}\hat{\boldsymbol{\delta}} = \left(\hat{\boldsymbol{x}} - \boldsymbol{x}^0\right)$ vector de correcções aos parâmetros desconhecidos

$_{n_1 \times 1}\hat{\boldsymbol{v}}_1 = \hat{\boldsymbol{\ell}}_1 - \boldsymbol{\ell}_1$ vector de resíduos do primeiro conjunto de observações

$_{n_2 \times 1}\hat{\boldsymbol{v}}_2 = \hat{\boldsymbol{\ell}}_2 - \boldsymbol{\ell}_2$ vector de resíduos do segundo conjunto de observações

$_{m_1 \times 1}\boldsymbol{w}_1 = \mathbf{f}_1\left(\boldsymbol{x}^0, \boldsymbol{\ell}_1\right)$ vector de fecho, relativo ao primeiro conjunto de observações

$_{m_2 \times 1}\boldsymbol{w}_2 = \mathbf{f}_2\left(\boldsymbol{x}^0, \boldsymbol{\ell}_2\right)$ vector de fecho, relativo ao segundo conjunto de observações

De novo se verifica que o modelo matemático resultante pode ser obtido a partir do modelo matemático do ajustamento combinado, recorrendo à partição das matrizes e vectores (com excepção do vector dos parâmetros):

$$\boxed{\mathbf{A}\hat{\boldsymbol{\delta}} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0}},$$

onde

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_2 \end{bmatrix}$$

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 \\ \hdashline \boldsymbol{w}_2 \end{bmatrix}$$

$$\hat{\boldsymbol{v}} = \begin{bmatrix} \hat{\boldsymbol{v}}_1 \\ \hdashline \hat{\boldsymbol{v}}_2 \end{bmatrix}$$

A matriz de variâncias-covariâncias será dada por

$$\mathbf{C}_\ell = \left[\begin{array}{c|c} \mathbf{C}_{\ell_1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{C}_{\ell_2} \end{array}\right].$$

Pretende-se agora minimizar a função

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}} \mathbf{P}_\ell \hat{\boldsymbol{v}}\,,$$

que explicitamente corresponde a

$$\phi = \left[\begin{array}{c|c} \hat{\boldsymbol{v}}_1^{\mathrm{T}} & \hat{\boldsymbol{v}}_2^{\mathrm{T}} \end{array}\right] \left[\begin{array}{c|c} \mathbf{P}_{\ell_1} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P}_{\ell_2} \end{array}\right] \left[\begin{array}{c} \hat{\boldsymbol{v}}_1 \\ \hline \hat{\boldsymbol{v}}_2 \end{array}\right] = \hat{\boldsymbol{v}}_1^{\mathrm{T}} \mathbf{P}_{\ell_1} \hat{\boldsymbol{v}}_1 + \hat{\boldsymbol{v}}_2^{\mathrm{T}} \mathbf{P}_{\ell_2} \hat{\boldsymbol{v}}_2\,,$$

sujeita aos constrangimentos

$$\mathbf{A}_1 \hat{\boldsymbol{\delta}} + \mathbf{B}_1 \hat{\boldsymbol{v}}_1 + \boldsymbol{w}_1 = \mathbf{0}$$

$$\mathbf{A}_2 \hat{\boldsymbol{\delta}} + \mathbf{B}_2 \hat{\boldsymbol{v}}_2 + \boldsymbol{w}_2 = \mathbf{0}\,.$$

Assim, a função de variação a minimizar é dada por:

$$\phi = \hat{\boldsymbol{v}}_1^{\mathrm{T}} \mathbf{P}_{\ell_1} \hat{\boldsymbol{v}}_1 + \hat{\boldsymbol{v}}_2^{\mathrm{T}} \mathbf{P}_{\ell_2} \hat{\boldsymbol{v}} + 2\hat{\mathbf{k}}_1^{\mathrm{T}} \left(\mathbf{A}_1 \hat{\boldsymbol{\delta}} + \mathbf{B}_1 \hat{\boldsymbol{v}}_1 + \boldsymbol{w}_1 \right) + 2\hat{\mathbf{k}}_2^{\mathrm{T}} \left(\mathbf{A}_2 \hat{\boldsymbol{\delta}} + \mathbf{B}_2 \hat{\boldsymbol{v}}_2 + \boldsymbol{w}_2 \right).$$

Calculando as diversas derivadas parciais, tem-se:

$$\frac{\partial \phi}{\partial \hat{\boldsymbol{v}}_1} = 2\hat{\boldsymbol{v}}_1^{\mathrm{T}} \mathbf{P}_{\ell_1} + 2\hat{\mathbf{k}}_1^{\mathrm{T}} \mathbf{B}_1 = \mathbf{0}\,,$$

$$\boxed{\mathbf{P}_{\ell_1} \hat{\boldsymbol{v}}_1 + \mathbf{B}_1^{\mathrm{T}} \hat{\mathbf{k}}_1 = \mathbf{0}}. \tag{3.23}$$

$$\frac{\partial \phi}{\partial \hat{\boldsymbol{v}}_2} = 2\hat{\boldsymbol{v}}_2^{\mathrm{T}} \mathbf{P}_{\ell_2} + 2\hat{\mathbf{k}}_2^{\mathrm{T}} \mathbf{B}_2 = \mathbf{0}\,,$$

$$\boxed{\mathbf{P}_{\ell_2} \hat{\boldsymbol{v}}_2 + \mathbf{B}_2^{\mathrm{T}} \hat{\mathbf{k}}_2 = \mathbf{0}}. \tag{3.24}$$

$$\frac{\partial \phi}{\partial \hat{\boldsymbol{\delta}}} = 2\hat{\mathbf{k}}_1^{\mathrm{T}} \mathbf{A}_1 + 2\hat{\mathbf{k}}_2^{\mathrm{T}} \mathbf{A}_2 = \mathbf{0}\,,$$

$$\boxed{\mathbf{A}_1^{\mathrm{T}} \hat{\mathbf{k}}_1 + \mathbf{A}_2^{\mathrm{T}} \hat{\mathbf{k}}_2 = \mathbf{0}}. \tag{3.25}$$

O sistema de equações a resolver é dado por

$$\begin{cases} \mathbf{A}_1\hat{\delta} + \mathbf{B}_1\hat{v}_1 + w_1 = \mathbf{0} \\ \mathbf{A}_2\hat{\delta} + \mathbf{B}_2\hat{v}_2 + w_2 = \mathbf{0} \\ \mathbf{P}_{\ell_1}\hat{v}_1 + \mathbf{B}_1^{\mathrm{T}}\hat{\mathbf{k}}_1 = \mathbf{0} \\ \mathbf{P}_{\ell_2}\hat{v}_2 + \mathbf{B}_2^{\mathrm{T}}\hat{\mathbf{k}}_2 = \mathbf{0} \\ \mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{k}}_1 + \mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2 = \mathbf{0} \end{cases} \tag{3.26}$$

que pode ser escrito matricialmente sob a forma

$$\left[\begin{array}{c|cccc} \mathbf{P}_{\ell_1} & \mathbf{0} & \mathbf{B}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \hline \mathbf{0} & \mathbf{P}_{\ell_2} & \mathbf{0} & \mathbf{B}_2^{\mathrm{T}} & \mathbf{0} \\ \mathbf{B}_1 & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_1 \\ \mathbf{0} & \mathbf{B}_2 & \mathbf{0} & \mathbf{0} & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_1^{\mathrm{T}} & \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v}_1 \\ \hline \hat{v}_2 \\ \hat{\mathbf{k}}_1 \\ \hat{\mathbf{k}}_2 \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hline \mathbf{0} \\ w_1 \\ w_2 \\ \mathbf{0} \end{array}\right] = \mathbf{0}\,. \tag{3.27}$$

A partição estabelecida permite eliminar $\hat{v}_1$, e o novo sistema de equações será dado por

$$\left[\begin{array}{c|ccc} \mathbf{P}_{\ell_2} & \mathbf{0} & \mathbf{B}_2^{\mathrm{T}} & \mathbf{0} \\ \hline \mathbf{0} & -\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{A}_1 \\ \mathbf{B}_2 & \mathbf{0} & \mathbf{0} & \mathbf{A}_2 \\ \mathbf{0} & \mathbf{A}_1^{\mathrm{T}} & \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v}_2 \\ \hline \hat{\mathbf{k}}_1 \\ \hat{\mathbf{k}}_2 \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hline w_1 \\ w_2 \\ \mathbf{0} \end{array}\right] = \mathbf{0}\,, \tag{3.28}$$

o que nos permite eliminar $\hat{v}_2$. Obtemos

$$\left[\begin{array}{c|cc} -\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{A}_1 \\ \hline \mathbf{0} & -\mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}} & \mathbf{A}_2 \\ \mathbf{A}_1^{\mathrm{T}} & \mathbf{A}_2^{\mathrm{T}} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}}_1 \\ \hline \hat{\mathbf{k}}_2 \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w_1 \\ \hline w_2 \\ \mathbf{0} \end{array}\right] = \mathbf{0}\,. \tag{3.29}$$

Deste sistema de equações eliminamos $\hat{\mathbf{k}}_1$ e obtemos

$$\left[\begin{array}{c|c} -\mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}} & \mathbf{A}_2 \\ \hline \mathbf{A}_2^{\mathrm{T}} & \mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}\right)^{-1}\mathbf{A}_1 \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}}_2 \\ \hline \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w_2 \\ \hline \mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}\right)^{-1} w_1 \end{array}\right] = \mathbf{0}\,. \tag{3.30}$$

Finalmente eliminamos $\hat{\mathbf{k}}_2$. Usando $\mathbf{M}_1 = \mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}$ e $\mathbf{M}_2 = \mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}}$, obtemos:

$$\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\mathbf{A}_2\right)\hat{\delta} + \mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}w_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}w_2 = \mathbf{0}$$

donde

$$\boxed{\hat{\delta} = -\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\mathbf{A}_2\right)^{-1}\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\boldsymbol{w}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\boldsymbol{w}_2\right)}. \tag{3.31}$$

O vector dos parâmetros desconhecidos será finalmente dado por

$$\boxed{\hat{\boldsymbol{x}} = \boldsymbol{x}^0 + \hat{\delta}}.$$

A análise da Equação (3.31) permite verificar que o vector de correcções é constituído pela soma de equações normais correspondentes a cada uma das épocas de observação, daí o este método ser também conhecido pelo **método da soma das equações normais**. O processo pode ser ampliado a qualquer número de épocas, ou seja,

$$\boxed{\hat{\delta} = -\left(\sum_{i=1}^{n}\mathbf{A}_i^{\mathrm{T}}\mathbf{M}_i^{-1}\mathbf{A}_i\right)^{-1}\left(\sum_{i=1}^{n}\mathbf{A}_i^{\mathrm{T}}\mathbf{M}_i^{-1}\boldsymbol{w}_i\right)}. \tag{3.32}$$

O método permite também determinar, de forma independente, os vectores de observações ajustadas para cada uma das épocas de observações.

Para determinar os resíduos ajustados para os dois conjuntos de observações, começamos por determinar os vectores de multiplicadores de Lagrange. Das primeiras equações de (3.30) e (3.29), obtemos respectivamente:

$$\boxed{\hat{\mathbf{k}}_2 = \left(\mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}_2\hat{\delta} + \boldsymbol{w}_2\right)} \tag{3.33}$$

e

$$\boxed{\hat{\mathbf{k}}_1 = \left(\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}_1\hat{\delta} + \boldsymbol{w}_1\right)}. \tag{3.34}$$

Os resíduos ajustados obtêm-se das equações (4) e (3) do sistema (3.26), ou seja,

$$\boxed{\hat{\boldsymbol{v}}_2 = -\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}}\hat{\mathbf{k}}_2} \tag{3.35}$$

$$\boxed{\hat{\boldsymbol{v}}_1 = -\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}\hat{\mathbf{k}}_1}, \tag{3.36}$$

o que permite determinar os dois vectores de observações ajustadas

$$\boxed{\hat{\boldsymbol{\ell}}_1 = \boldsymbol{\ell}_1 - \mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^{\mathrm{T}}\hat{\mathbf{k}}_1} \tag{3.37}$$

$$\boxed{\hat{\boldsymbol{\ell}}_2 = \boldsymbol{\ell}_2 - \mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^{\mathrm{T}}\hat{\mathbf{k}}_2}. \tag{3.38}$$

Matriz de variâncias-covariâncias dos parâmetros

Para deduzir a expressão para a matriz de variâncias-covariâncias, comecemos por escrever a expressão (3.31) sob a forma:

$$\hat{\boldsymbol{\delta}} = -\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\boldsymbol{w}_1 - \left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\boldsymbol{w}_2$$

A matriz de variâncias-covariâncias para os parâmetros ajustados obtém-se aplicando a lei de propagação de variâncias-covariâncias

$$\mathbf{C}_{\hat{\delta}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}} = \left[\begin{array}{c:c} \dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_1} & \dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_2} \end{array}\right]\left[\begin{array}{c:c} \sigma_0^2\mathbf{P}_{\ell_1}^{-1} & 0 \\ \hdashline 0 & \sigma_0^2\mathbf{P}_{\ell_2}^{-1} \end{array}\right]\left[\begin{array}{c} \left(\dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_1}\right)^{\mathrm{T}} \\ \hdashline \left(\dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_2}\right)^{\mathrm{T}} \end{array}\right],$$

ou seja,

$$\mathbf{C}_{\hat{\delta}} = \sigma_0^2\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_1}\right)\mathbf{P}_{\ell_1}^{-1}\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_1}\right)^{\mathrm{T}} + \sigma_0^2\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_2}\right)\mathbf{P}_{\ell_2}^{-1}\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_2}\right)^{\mathrm{T}}.$$

Dado que

$$\boldsymbol{w}_1 = \mathbf{f}_1\left(\boldsymbol{x}^0, \boldsymbol{\ell}_1\right)$$

e

$$\boldsymbol{w}_2 = \mathbf{f}_2\left(\boldsymbol{x}^0, \boldsymbol{\ell}_2\right)$$

tem-se:

$$\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}_1} = -\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^{\mathrm{T}}\mathbf{M}_2^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}_1^{-1}\mathbf{B}_1$$

e

$$\frac{\partial\hat{\delta}}{\partial\boldsymbol{\ell}_2} = -\left(\mathbf{A}_1^T\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^T\mathbf{M}_2^{-1}\mathbf{A}_2\right)^{-1}\mathbf{A}_2^T\mathbf{M}_2^{-1}\mathbf{B}_2 .$$

Seja **Z** a matriz auxiliar seguinte:

$$\mathbf{Z} = \left(\mathbf{A}_1^T\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^T\mathbf{M}_2^{-1}\mathbf{A}_2\right).$$

Nestas condições, tem-se

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\mathbf{Z}^{-1}\mathbf{A}_1^T\mathbf{M}_1^{-1}\underbrace{\left(\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^T\right)}_{\mathbf{M}_1}\mathbf{M}_1^{-1}\mathbf{A}_1\mathbf{Z}^{-1} + \sigma_0^2\mathbf{Z}^{-1}\mathbf{A}_2^T\mathbf{M}_2^{-1}\underbrace{\left(\mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^T\right)}_{\mathbf{M}_2}\mathbf{M}_2^{-1}\mathbf{A}_2\mathbf{Z}^{-1}$$

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\mathbf{Z}^{-1}\mathbf{A}_1^T\mathbf{M}_1^{-1}\mathbf{A}_1\mathbf{Z}^{-1} + \mathbf{Z}^{-1}\mathbf{A}_2^T\mathbf{M}_2^{-1}\mathbf{A}_2\mathbf{Z}^{-1}\right)$$

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\mathbf{Z}^{-1}\underbrace{\left(\mathbf{A}_1^T\mathbf{M}_1^{-1}\mathbf{A}_1 + \mathbf{A}_2^T\mathbf{M}_2^{-1}\mathbf{A}_2\right)}_{\mathbf{Z}}\mathbf{Z}^{-1}$$

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\mathbf{A}_1^T\left(\mathbf{B}_1\mathbf{P}_{\ell_1}^{-1}\mathbf{B}_1^T\right)^{-1}\mathbf{A}_1 + \mathbf{A}_2^T\left(\mathbf{B}_2\mathbf{P}_{\ell_2}^{-1}\mathbf{B}_2^T\right)^{-1}\mathbf{A}_2\right)^{-1}}. \tag{3.39}$$

À semelhança do vector de correcções aos parâmetros, também a matriz de variâncias-covariâncias é baseada na soma de equações normais, pelo que pode ser generalizada para k conjuntos de observações:

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\sum_{i=1}^{k}\mathbf{A}_i^T\left(\mathbf{B}_i\mathbf{P}_{\ell_i}^{-1}\mathbf{B}_i^T\right)^{-1}\mathbf{A}_i\right)^{-1}}. \tag{3.40}$$

Generalizando, obtemos para variância *a posteriori*

$$\boxed{\hat{\sigma}_0^2 = \frac{\sum_{i=1}^{k}\hat{\boldsymbol{v}}_i^T\mathbf{P}_{\ell_i}\hat{\boldsymbol{v}}_i}{\mathrm{df}}}, \tag{3.41}$$

onde df=n-u, com $n = \sum_{i=1}^{k} n_i$.

Ajustamento com Restrições entre Parâmetros

Vamos supor que existe um conhecimento adicional de relações entre os parâmetros desconhecidos, que não é expresso pelo modelo matemático que relaciona as observações e os parâmetros desconhecidos (por exemplo, sabe-se que a soma de dois ângulos é de 180°, ou que as coordenadas de um satélite se ajustam a uma trajectória elíptica). Se estas relações funcionais adicionais forem descritas pelo modelo matemático

$$\mathbf{f}_2(\hat{x}) = \mathbf{0}\,,$$

o modelo matemático completo será dado por

$$\begin{cases} \mathbf{f}_1\left(\hat{x}, \hat{\ell}\right) = \mathbf{0} \\ \mathbf{f}_2\left(\hat{x}\right) = \mathbf{0} \end{cases}$$

Linearizando ambas as funções em torno de valores aproximados x^0 e dos valores observados ℓ, vem:

$$\mathbf{f}_1\left(\hat{x}, \hat{\ell}\right) \doteq \mathbf{f}_1\left(x^0, \ell\right) + \left.\frac{\partial \mathbf{f}_1}{\partial \hat{\mathbf{x}}}\right|_{x^0, \ell} \left(\hat{x} - x^0\right) + \left.\frac{\partial \mathbf{f}_1}{\partial \hat{\ell}}\right|_{x^0, \ell} \left(\hat{\ell} - \ell\right) = \mathbf{0}$$

$$\mathbf{f}_2\left(\hat{x}\right) \doteq \mathbf{f}_2\left(x^0\right) + \left.\frac{\partial \mathbf{f}_2}{\partial \hat{\mathbf{x}}}\right|_{x^0} \left(\hat{x} - x^0\right) = \mathbf{0}$$

ou, em forma matricial:

$$\boxed{\mathbf{A}_1\hat{\delta} + \mathbf{B}\hat{v} + w_1 = \mathbf{0}}, \tag{3.42}$$

$$\boxed{\mathbf{A}_2\hat{\delta} + w_2 = \mathbf{0}}, \tag{3.43}$$

onde

$${}_{m_1 \times u}\mathbf{A}_1 = \left.\frac{\partial \mathbf{f}_1}{\partial \hat{x}}\right|_{x^0, \ell}$$ primeira matriz de configuração

$${}_{m_2 \times u}\mathbf{A}_2 = \left.\frac{\partial \mathbf{f}_2}{\partial \hat{x}}\right|_{x^0, \ell}$$ primeira matriz de configuração, relativa à equação de constrangimento entre parâmetros

$_{m_1\times n}\mathbf{B} = \left.\dfrac{\partial \mathbf{f}_1}{\partial \hat{\boldsymbol{\ell}}}\right|_{x^0,\boldsymbol{\ell}}$ segunda matriz de configuração

$_{u\times 1}\hat{\delta} = \left(\hat{\boldsymbol{x}} - \boldsymbol{x}^0\right)$ vector de correcções aos parâmetros desconhecidos

$_{n\times 1}\hat{\boldsymbol{v}} = \hat{\boldsymbol{\ell}} - \boldsymbol{\ell}$ vector de resíduos

$_{m_1\times 1}\boldsymbol{w}_1 = \mathbf{f}_1\left(\boldsymbol{x}^0, \boldsymbol{\ell}\right)$ vector de fecho

$_{m_2\times 1}\boldsymbol{w}_2 = \mathbf{f}_2\left(\boldsymbol{x}^0\right)$ vector de fecho, relativo à equação de constrangimento.

Este novo modelo matemático pode ser derivado do caso geral do ajustamento combinado,

$$\mathbf{A}\hat{\delta} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w} = \mathbf{0},$$

com

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 \\ \hdashline \mathbf{A}_2 \end{bmatrix}$$

$$\mathbf{B} = \left[\begin{array}{c:c} \mathbf{B} & \mathbf{0} \\ \hdashline \mathbf{0} & \mathbf{0} \end{array}\right]$$

e

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_1 \\ \hdashline \boldsymbol{w}_2 \end{bmatrix}.$$

Pretendemos minimizar a função

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\hat{\boldsymbol{v}},$$

com as restrições:

$$\mathbf{A}_1\hat{\delta} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w}_1 = \mathbf{0},$$

$$\mathbf{A}_2\hat{\delta} + \boldsymbol{w}_2 = \mathbf{0}.$$

A função de variação será então dada por:

$$\phi = \hat{\boldsymbol{v}}^{\mathrm{T}}\mathbf{P}_{\boldsymbol{\ell}}\hat{\boldsymbol{v}} + 2\hat{\mathbf{k}}_1^{\mathrm{T}}\left(\mathbf{A}_1\hat{\delta} + \mathbf{B}\hat{\boldsymbol{v}} + \boldsymbol{w}_1\right) + 2\hat{\mathbf{k}}_2^{\mathrm{T}}\left(\mathbf{A}_2\hat{\delta} + \boldsymbol{w}_2\right).$$

Calculando as diversas derivadas parciais, tem-se:

$$\frac{\partial \phi}{\partial \hat{v}} = 2\hat{v}^{\mathrm{T}}\mathbf{P}_{\ell} + 2\hat{\mathbf{k}}_1^{\mathrm{T}}\mathbf{B} = \mathbf{0},$$

$$\boxed{\mathbf{P}_{\ell}\hat{v} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_1 = \mathbf{0}}. \tag{3.44}$$

$$\frac{\partial \phi}{\partial \hat{\delta}} = 2\hat{\mathbf{k}}_1^{\mathrm{T}}\mathbf{A}_1 + 2\hat{\mathbf{k}}_2^{\mathrm{T}}\mathbf{A}_2 = \mathbf{0},$$

$$\boxed{\mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{k}}_1 + \mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2 = \mathbf{0}}. \tag{3.45}$$

O sistema de equações a resolver é dado por

$$\begin{cases} \mathbf{A}_1\hat{\delta} + \mathbf{B}\hat{v} + w_1 = \mathbf{0} \\ \mathbf{A}_2\hat{\delta} + w_2 = \mathbf{0} \\ \mathbf{P}_{\ell}\hat{v} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_1 = \mathbf{0} \\ \mathbf{A}_1^{\mathrm{T}}\hat{\mathbf{k}}_1 + \mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2 = \mathbf{0} \end{cases} \tag{3.46}$$

que pode ser escrito matricialmente sob a forma

$$\left[\begin{array}{c:ccc} \mathbf{P}_{\ell} & \mathbf{B}^{\mathrm{T}} & \mathbf{0} & \mathbf{0} \\ \hdashline \mathbf{B} & \mathbf{0} & \mathbf{A}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{A}_2^{\mathrm{T}} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v} \\ \hdashline \hat{\mathbf{k}}_1 \\ \hat{\delta} \\ \hat{\mathbf{k}}_2 \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline w_1 \\ \mathbf{0} \\ w_2 \end{array}\right] = \mathbf{0}. \tag{3.47}$$

A partição estabelecida permite eliminar $\hat{v}$ e obter o sistema de equações seguinte:

$$\left[\begin{array}{c:cc} -\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}} & \mathbf{A}_1 & \mathbf{0} \\ \hdashline \mathbf{A}_1^{\mathrm{T}} & \mathbf{0} & \mathbf{A}_2^{\mathrm{T}} \\ \mathbf{0} & \mathbf{A}_2 & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}}_1 \\ \hdashline \hat{\delta} \\ \hat{\mathbf{k}}_2 \end{array}\right] + \left[\begin{array}{c} w_1 \\ \hdashline \mathbf{0} \\ w_2 \end{array}\right] = \mathbf{0}. \tag{3.48}$$

Deste sistema de equações eliminamos $\hat{\mathbf{k}}_1$, donde

$$\left[\begin{array}{c:c} \mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}_1 & \mathbf{A}_2^{\mathrm{T}} \\ \hdashline \mathbf{A}_2 & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\delta} \\ \hdashline \hat{\mathbf{k}}_2 \end{array}\right] + \left[\begin{array}{c} \mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_{\ell}^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1} w_1 \\ \hdashline w_2 \end{array}\right] = \mathbf{0}. \tag{3.49}$$

Eliminando $\hat{\delta}$ e usando $\mathbf{M} = \mathbf{B}\mathbf{P}_{\ell_1}^{-1}\mathbf{B}^{\mathrm{T}}$, obtemos

$$-\mathbf{A}_2\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2+\left(\boldsymbol{w}_2-\mathbf{A}_2\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1\right)=\mathbf{0}$$

ou

$$\boxed{\hat{\mathbf{k}}_2=\left(\mathbf{A}_2\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\left(\boldsymbol{w}_2-\mathbf{A}_2\left(\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\right)^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1\right)}. \qquad (3.50)$$

Retomando a primeira equação do sistema (3.49), obtemos:

$$\boxed{\hat{\delta}=-\left(\mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}_1\right)^{-1}\left(\mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2+\mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\boldsymbol{w}_1\right)}. \qquad (3.51)$$

O vector dos parâmetros desconhecidos será finalmente dado por

$$\boxed{\hat{x}=x^0+\hat{\delta}}.$$

Para determinar os resíduos ajustados obtemos os vectores de multiplicadores de Lagrange da primeira equação do sistema (3.48)

$$\boxed{\hat{\mathbf{k}}_1=\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}_1\hat{\delta}+\boldsymbol{w}_1\right)}. \qquad (3.52)$$

Da primeira equação do sistema (3.47), obtemos

$$\boxed{\hat{v}=-\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_1}, \qquad (3.53)$$

o que permite determinar o vector de observações ajustadas

$$\boxed{\hat{\ell}=\ell-\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}_1}. \qquad (3.54)$$

Matriz de variâncias-covariâncias dos parâmetros

Para determinar a matriz de variâncias-covariâncias das correcções aos parâmetros ajustados (e dos parâmetros ajustados propriamente ditos), comecemos por reescrever a Equação (3.51) sob a forma:

$$\hat{\delta}=-\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\hat{\mathbf{k}}_2-\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1,$$

onde

$$\mathbf{N}_{11}=\mathbf{A}_1^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}_1=\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1.$$

Por outro lado, a expressão de $\hat{\mathbf{k}}_2$ é

$$\hat{\mathbf{k}}_2 = \left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\boldsymbol{w}_2 - \left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1,$$

ou seja

$$\hat{\delta} = -\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\boldsymbol{w}_2 + \mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1 - \mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w}_1$$

A matriz de variâncias-covariâncias para os parâmetros ajustados obtém-se aplicando a lei de propagação de variâncias-covariâncias

$$\mathbf{C}_{\hat{\delta}} = \mathbf{J}\mathbf{C}_{\ell}\mathbf{J}^{\mathrm{T}}$$

Dado que $\boldsymbol{w}_2$ não depende de $\boldsymbol{\ell}$ e $\boldsymbol{w}_1 = \mathbf{f}_1\left(\boldsymbol{x}^0, \boldsymbol{\ell}\right)$, então

$$\frac{\partial\hat{\delta}}{\partial\ell} = \mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} - \mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B},$$

logo

$$\begin{aligned} \mathrm{C}_{\hat{\delta}} = \sigma_0^2 & \left(\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} - \mathbf{N}_{11}^{-1}\mathbf{A}_1^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}\right)\mathbf{P}_{\ell}^{-1} \\ & \left(\mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1} - \mathbf{B}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}_1\mathbf{N}_{11}^{-1}\right). \end{aligned}$$

Desenvolvendo e simplificando, obtemos

$$\mathrm{C}_{\hat{\delta}} = \sigma_0^2\left(\mathbf{N}_{11}^{-1} - \mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}\right)$$

ou

$$\boxed{\mathrm{C}_{\hat{\delta}} = \sigma_0^2\mathbf{N}_{11}^{-1} - \sigma_0^2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\left(\mathbf{A}_2\mathbf{N}_{11}^{-1}\mathbf{A}_2^{\mathrm{T}}\right)^{-1}\mathbf{A}_2\mathbf{N}_{11}^{-1}}. \qquad (3.55)$$

Ajustamento com Parâmetros Parcialmente Conhecidos

Em todos os casos analisados até aqui, considerámos sempre que os modelos matemáticos representavam de forma rigorosa as relações existentes entre os parâmetros desconhecidos e o conjunto de observações. Existem casos, no entanto, em que os parâmetros não são totalmente desconhecidos. Como exemplo, na densificação de redes geodésicas, que inclui vértices para os quais se conhecem as coordenadas de um ajustamento prévio, estas podem ser consideradas não como totalmente conhecidas (fixas no processo de ajustamento), mas parcialmente conhecidas. Neste caso, as coordenadas destes vértices deixarão de ter variâncias infinitas (peso nulo) e são consideradas como estimativas preliminares dos parâmetros, com uma variância associada; são, desta forma, tratadas como se fossem observações, pelo que as denominaremos por ***pseudo-observações***. Esta estratégia faz com que as estimativas resultantes para estes parâmetros venham a ser constrangidas pelas variâncias da estimativa preliminar.

O modelo matemático para o ajustamento com parâmetros parcialmente conhecidos, denominado na literatura anglo-saxónica como "*weighted constraints*", é semelhante ao caso do usado no modelo condicional, com as diferenças inerentes ao facto do vector de observações conter simultaneamente observações e parâmetros (pseudo-observações). Neste caso, os parâmetros iniciais já não são valores arbitrários, pelo que iremos designar esse vector inicial por $\boldsymbol{x}$. Tem-se:

$$\mathbf{f}\left(\hat{\boldsymbol{\ell}}^{*}\right)=\mathbf{0},$$

sendo

$$\boldsymbol{\ell}^{*}=\begin{bmatrix}\boldsymbol{\ell}\\ \hdashline \boldsymbol{x}\end{bmatrix},$$

com

$$\mathbf{C}_{\ell^{*}}=\left[\begin{array}{c:c}\mathbf{C}_{\ell} & \mathbf{0}\\ \hdashline \mathbf{0} & \mathbf{C}_{x}\end{array}\right],$$

$$\mathbf{P}_{\ell^*} = \left[\begin{array}{c:c} \mathbf{P}_\ell & \mathbf{0} \\ \hdashline \mathbf{0} & \mathbf{P}_x \end{array}\right],$$

onde $\mathbf{P}_\ell = \sigma_0^2 \mathbf{C}_\ell^{-1}$ e $\mathbf{P}_x = \sigma_0^2 \mathbf{C}_x^{-1}$.

Linearizando, obtemos

$$\mathbf{f}_1\left(\hat{\ell}^*\right) \doteq \mathbf{f}_1\left(x, \ell^*\right) + \left.\frac{\partial \mathbf{f}_1}{\partial \hat{\ell}^*}\right|_{x,\ell^*} \left(\hat{\ell}^* - \ell^*\right) = \mathbf{0}$$

que é equivalente a

$$\mathbf{f}\left(\hat{\ell}^*\right) \doteq \mathbf{f}\left(\ell^*\right) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\ell}}\right|_{\ell} \left(\hat{\ell} - \ell\right) + \left.\frac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}}\right|_{\ell} \left(\hat{x} - x\right) = \mathbf{0}.$$

Assim, a forma linearizada do modelo matemático é

$$\boxed{\mathbf{B}^* v^* + w = \mathbf{0}}, \tag{4.48}$$

onde

$$\hat{v}^* = \left[\begin{array}{c} \hat{v} \\ \hdashline \hat{\delta} \end{array}\right]$$

$$\mathbf{B}^* = \left[\begin{array}{c:c} \mathbf{B} & \mathbf{A} \end{array}\right]$$

com

$_{m\times u}\mathbf{A} = \left.\dfrac{\partial \mathbf{f}}{\partial \hat{\mathbf{x}}}\right\|_{x,\ell}$	primeira matriz de configuração
$_{m\times n}\mathbf{B} = \left.\dfrac{\partial \mathbf{f}}{\partial \hat{\ell}}\right\|_{x,\ell}$	segunda matriz de configuração
$_{u\times 1}\hat{\delta} = \left(\hat{x} - x\right)$	vector de correcções aos parâmetros parcialmente conhecidos (pseudo-observações)
$_{n\times 1}\hat{v} = \hat{\ell} - \ell$	vector de resíduos
$_{m\times 1}w = \mathbf{f}\left(\ell^*\right)$	vector de fecho

Pretendemos minimizar a função

$$\phi = \hat{v}^{*\mathrm{T}} \mathbf{P}_{\ell} \hat{v}^{*} ,$$

ou seja,

$$\phi = \begin{bmatrix} \hat{v} \\ \text{---} \\ \hat{\delta} \end{bmatrix}^{\mathrm{T}} \mathbf{P}_{\ell}^{*} \begin{bmatrix} \hat{v} \\ \text{---} \\ \hat{\delta} \end{bmatrix} = \hat{v}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{v} + \hat{\delta}^{\mathrm{T}} \mathbf{P}_{x} \hat{\delta} ,$$

com o constrangimento

$$\mathbf{A}\hat{\delta} + \mathbf{B}\hat{v} + w = \mathbf{0} .$$

A função de variação será então dada por

$$\phi = \hat{v}^{\mathrm{T}} \mathbf{P}_{\ell} \hat{v} + \hat{\delta}^{\mathrm{T}} \mathbf{P}_{x} \hat{\delta} + 2\hat{\mathbf{k}}^{\mathrm{T}} \left(\mathbf{A}\hat{\delta} + \mathbf{B}\hat{v} + w \right).$$

Calculando as diversas derivadas parciais, tem-se:

$$\frac{\partial \phi}{\partial \hat{v}} = 2\hat{v}^{\mathrm{T}} \mathbf{P}_{\ell} + 2\hat{\mathbf{k}}^{\mathrm{T}} \mathbf{B} = \mathbf{0} ,$$

$$\boxed{\mathbf{P}_{\ell}\hat{v} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}} . \tag{3.56}$$

$$\frac{\partial \phi}{\partial \hat{\delta}} = 2\hat{\delta}^{\mathrm{T}} \mathbf{P}_{x} + 2\hat{\mathbf{k}}^{\mathrm{T}} \mathbf{A} = \mathbf{0} ,$$

$$\boxed{\mathbf{P}_{x}\hat{\delta} + \mathbf{A}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0}} . \tag{3.57}$$

O sistema de equações a resolver é dado por

$$\begin{cases} \mathbf{A}\hat{\delta} + \mathbf{B}\hat{v} + w = \mathbf{0} \\ \mathbf{P}_{\ell}\hat{v} + \mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \\ \mathbf{P}_{x}\hat{\delta} + \mathbf{A}^{\mathrm{T}}\hat{\mathbf{k}} = \mathbf{0} \end{cases} \tag{3.58}$$

que pode ser reorganizado e escrito matricialmente sob a forma

$$\left[\begin{array}{c:cc} \mathbf{P}_{\ell} & \mathbf{B}^{\mathrm{T}} & \mathbf{0} \\ \hdashline \mathbf{B} & \mathbf{0} & \mathbf{A} \\ \mathbf{0} & \mathbf{A}^{\mathrm{T}} & \mathbf{P}_{x} \end{array}\right] \begin{bmatrix} \hat{v} \\ \text{---} \\ \hat{\mathbf{k}} \\ \hat{\delta} \end{bmatrix} + \begin{bmatrix} \mathbf{0} \\ \text{----} \\ w \\ \mathbf{0} \end{bmatrix} = \mathbf{0} . \tag{3.59}$$

A partição estabelecida permite eliminar $\hat{v}$ e obter o sistema de equações seguinte:

$$\left[\begin{array}{c|c} -\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}} & \mathbf{A} \\ \hline \mathbf{A}^{\mathrm{T}} & \mathbf{P}_x \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}} \\ \hline \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w \\ \hline \mathbf{0} \end{array}\right] = \mathbf{0}\,. \tag{3.60}$$

Deste sistema de equações eliminamos $\hat{\mathbf{k}}$ e obtemos

$$\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)\hat{\delta} + \mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1} w = \mathbf{0}\,,$$

donde

$$\boxed{\hat{\delta} = -\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1} w} \tag{3.61}$$

e

$$\boxed{\hat{x} = x + \hat{\delta}}.$$

Da primeira equação de (3.60) obtemos

$$\boxed{\hat{\mathbf{k}} = \left(\mathbf{B}\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\right)^{-1}\left(\mathbf{A}\hat{\delta} + w\right)}. \tag{3.62}$$

Utilizando (3.59), obtemos o vector dos resíduos ajustados:

$$\boxed{\hat{v} = -\mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}}, \tag{3.63}$$

donde

$$\boxed{\hat{\ell} = \ell - \mathbf{P}_\ell^{-1}\mathbf{B}^{\mathrm{T}}\hat{\mathbf{k}}}. \tag{3.64}$$

Matrizes de variâncias-covariâncias

Dado que os parâmetros são estimados a partir de estimativas prévias com variâncias associadas, o procedimento para o cálculo da matriz de variâncias-covariâncias para os parâmetros ajustados é diferente, dado que $\mathbf{C}_{\hat{x}} \neq \mathbf{C}_{\hat{\delta}}$. Com efeito, para o vector de correcções, tem-se

$$\hat{\delta} = -\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1} w\,.$$

Como $\boldsymbol{x}$ é um vector de pseudo-observações e $\boldsymbol{w} = \mathbf{f}(\boldsymbol{x}, \boldsymbol{\ell})$, vem

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \mathbf{J}\mathbf{C}_{\boldsymbol{\ell}^*}\mathbf{J}^{\mathrm{T}} = \left[\begin{array}{c:c} \dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{x}} & \dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}} \end{array}\right] \left[\begin{array}{c:c} \sigma_0^2\mathbf{P}_x^{-1} & 0 \\ \hdashline 0 & \sigma_0^2\mathbf{P}_{\boldsymbol{\ell}}^{-1} \end{array}\right] \left[\begin{array}{c} \left(\dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{x}}\right)^{\mathrm{T}} \\ \hdashline \left(\dfrac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}}\right)^{\mathrm{T}} \end{array}\right]$$

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{x}}\right)\mathbf{P}_x^{-1}\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{x}}\right)^{\mathrm{T}} + \sigma_0^2\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}}\right)\mathbf{P}_{\boldsymbol{\ell}}^{-1}\left(\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}}\right)^{\mathrm{T}}$$

sendo

$$\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{\ell}} = -\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} = -\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B}$$

e

$$\frac{\partial\hat{\boldsymbol{\delta}}}{\partial\boldsymbol{x}} = -\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A} = -\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N} \ ,$$

com

$$\mathbf{N} = \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A} \, .$$

Assim, tem-se

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N}\mathbf{P}_x^{-1}\mathbf{N}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1} + \sigma_0^2\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\underbrace{\mathbf{B}\mathbf{P}_{\boldsymbol{\ell}}^{-1}\mathbf{B}^{\mathrm{T}}}_{\mathbf{M}}\mathbf{M}^{-1}\mathbf{A}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}$$

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}} = \sigma_0^2\left(\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N}\mathbf{P}_x^{-1}\mathbf{N}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1} + \left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\right)} \qquad (3.65)$$

Por outro lado, tem-se

$$\hat{\boldsymbol{x}} = \boldsymbol{x} - \left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\boldsymbol{w} \, ,$$

pelo que

$$\frac{\partial\hat{\boldsymbol{x}}}{\partial\boldsymbol{x}} = \mathbf{I} - \left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A} = \mathbf{I} - \left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N}$$

e

$$\frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{\ell}} = -\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} = -\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{B} \ .$$

Aplicando a lei de propagação de variâncias-covariâncias, tem-se

$$\mathbf{C}_{\hat{x}} = \sigma_0^2\left(\frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{x}}\right)\mathbf{P}_x^{-1}\left(\frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{x}}\right)^{\mathrm{T}} + \sigma_0^2\left(\frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{\ell}}\right)\mathbf{P}_{\ell}^{-1}\left(\frac{\partial \hat{\boldsymbol{x}}}{\partial \boldsymbol{\ell}}\right)^{\mathrm{T}}$$

ou seja

$$\mathbf{C}_{\hat{x}} = \sigma_0^2\left(\mathbf{I} - \left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{N}\right)\mathbf{P}_x^{-1}\left(\mathbf{I} - \mathbf{N}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\right) + \sigma_0^2\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}\mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}.$$

Colocando $\left(\mathbf{P}_x + \mathbf{N}\right)^{-1}$ em evidência e simplificando, obtemos

$$\boxed{\mathbf{C}_{\hat{x}} = \sigma_0^2\left(\mathbf{P}_x + \mathbf{A}^{\mathrm{T}}\mathbf{M}^{-1}\mathbf{A}\right)^{-1}}. \qquad (3.66)$$

Facilmente se reconhece uma grande semelhança com a matriz correspondente num ajustamento combinado, que se obtém da Equação (3.66) fazendo $\mathbf{P}_x = \mathbf{0}$.

4. AJUSTAMENTOS PASSO A PASSO

Nos métodos estudados anteriormente, as observações são sempre ajustadas globalmente, o que se revela pouco conveniente nos casos em que o número de observações é demasiado grande, devido aos requisitos de memória e esforço computacional que são exigidos. Para contornar este tipo de problemas, existem os chamados métodos de mínimos quadrados **passo a passo** (*step-by-step*), entre os quais se incluem o ajustamento sequencial, o ajustamento de Bayes e o ajustamento por filtro de Kalman, por exemplo. Estes dois últimos são equivalentes matematicamente, mas diferentes sob o ponto de vista computacional. O ajustamento sequencial é usado em situações em que o número total de observações não está disponível num dado instante (ou se pretende ajustar um número elevado de observações por fases). À medida que essas observações se tornam disponíveis, as estimativas dos parâmetros determinadas até esse instante são actualizadas. O ajustamento por filtro de Kalman é utilizado nos casos em que os próprios parâmetros a estimar variam no tempo, como acontece com as técnicas de posicionamento cinemático.

Ajustamento Sequencial

O ajustamento sequencial é particularmente útil nos casos em que se têm muitas séries de observações, relativas a um determinado conjunto de parâmetros a estimar, obtidas em épocas diferentes. Contrariamente ao que se passa com o ajustamento com observações adicionais, em que todas as observações são tratadas em bloco, o ajustamento sequencial estima correcções às estimativas obtidas utilizando as observações da série anterior, facilitando o cálculo e as exigências de memória computacional, dado que as observações utilizadas nos passos anteriores já não serão incluídas.

Consideremos duas séries de observações obtidas em duas épocas distintas (t_{k-1} e $t_{k,}$ às quais associamos os índices k e k-1, respectivamente). O modelo matemático é do tipo:

$$\mathbf{f}_{k-1}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_{k-1}\right)=\mathbf{0}$$

$$\mathbf{f}_{k}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_{k}\right)=\mathbf{0}\,,$$

onde $\hat{\boldsymbol{\ell}}_{k-1}$ designa o vector de observações para a época t_{k-1}, de dimensões $n_1 \times 1$, e $\hat{\boldsymbol{\ell}}_k$ o vector de observações para a época t_k, de dimensões $n_2 \times 1$, e aos quais associamos as matrizes de variâncias-covariâncias $\mathbf{C}_{\boldsymbol{\ell}_{k-1}}$ e $\mathbf{C}_{\boldsymbol{\ell}_k}$, respectivamente.

Linearizando as duas funções usando séries de Taylor, em torno de valores aproximados $\boldsymbol{x}^0$ e dos valores observados $\boldsymbol{\ell}_{k-1}$ e $\boldsymbol{\ell}_k$, obtemos:

$$\mathbf{f}_{k-1}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_{k-1}\right)\doteq\mathbf{f}_{k-1}\left(\boldsymbol{x}^0,\boldsymbol{\ell}_{k-1}\right)+\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k-1}}\left(\hat{\boldsymbol{x}}-\boldsymbol{x}^0\right)+\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\hat{\boldsymbol{\ell}}_{k-1}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k-1}}\left(\hat{\boldsymbol{\ell}}_{k-1}-\boldsymbol{\ell}_{k-1}\right)=\mathbf{0}$$

$$\mathbf{f}_{k}\left(\hat{\boldsymbol{x}},\hat{\boldsymbol{\ell}}_{k}\right)\doteq\mathbf{f}_{k}\left(\boldsymbol{x}^0,\boldsymbol{\ell}_{k}\right)+\left.\frac{\partial\mathbf{f}_{k}}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k}}\left(\hat{\boldsymbol{x}}-\boldsymbol{x}^0\right)+\left.\frac{\partial\mathbf{f}_{k}}{\partial\hat{\boldsymbol{\ell}}_{k}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k}}\left(\hat{\boldsymbol{\ell}}_{k}-\boldsymbol{\ell}_{k}\right)=\mathbf{0}$$

ou, em forma matricial:

$$\boxed{\mathbf{A}_{k-1}\boldsymbol{\delta}+\mathbf{B}_{k-1}\boldsymbol{v}_{k-1}+\boldsymbol{w}_{k-1}=\mathbf{0}} \tag{4.1}$$

$$\boxed{\mathbf{A}_{k}\boldsymbol{\delta}+\mathbf{B}_{k}\boldsymbol{v}_{k}+\boldsymbol{w}_{k}=\mathbf{0}}\,, \tag{4.2}$$

onde

$${}_{m_1\times u}\mathbf{A}_{k-1}=\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k-1}}$$ primeira matriz de configuração, relativa à época t_{k-1}

$${}_{m_2\times u}\mathbf{A}_{k}=\left.\frac{\partial\mathbf{f}_{k}}{\partial\hat{\boldsymbol{x}}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k}}$$ primeira matriz de configuração, relativa à época t_k

$${}_{m_1\times n_1}\mathbf{B}_{k-1}=\left.\frac{\partial\mathbf{f}_{k-1}}{\partial\hat{\boldsymbol{\ell}}_{k-1}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k-1}}$$ segunda matriz de configuração, relativa à época t_{k-1}

$${}_{m_2\times n_2}\mathbf{B}_{k}=\left.\frac{\partial\mathbf{f}_{k}}{\partial\hat{\boldsymbol{\ell}}_{k}}\right|_{\boldsymbol{x}^0,\boldsymbol{\ell}_{k}}$$ segunda matriz de configuração, relativa à época t_k

$${}_{u\times 1}\boldsymbol{\delta}=\left(\hat{\boldsymbol{x}}-\boldsymbol{x}^0\right)$$ vector de correcções aos parâmetros desconhecidos

$${}_{n_1\times 1}\boldsymbol{v}_{k-1}=\hat{\boldsymbol{\ell}}_{k-1}-\boldsymbol{\ell}_{k-1}$$ vector de resíduos das observações da época t_{k-1}

$_{n_2 \times 1}\boldsymbol{v}_k = \hat{\boldsymbol{\ell}}_k - \boldsymbol{\ell}_k$ vector de resíduos das observações da época t_k

$_{m_1 \times 1}\boldsymbol{w}_{k-1} = \mathbf{f}\left(\boldsymbol{x}^0, \boldsymbol{\ell}_{k-1}\right)$ vector de fecho, relativo à época t_{k-1}

$_{m_2 \times 1}\boldsymbol{w}_k = \mathbf{f}\left(\boldsymbol{x}^0, \boldsymbol{\ell}_k\right)$ vector de fecho, relativo à época t_k

Os passos a seguir são idênticos aos usados no caso de um ajustamento com observações adicionais. O modelo matemático resultante pode obter-se do modelo matemático do ajustamento combinado, recorrendo à partição das matrizes e vectores:

$$\boxed{\mathbf{A}\delta + \mathbf{B}\boldsymbol{v} + \boldsymbol{w} = \mathbf{0}},$$

onde

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_{k-1} \\ \hdashline \mathbf{A}_k \end{bmatrix}$$

$$\mathbf{B} = \begin{bmatrix} \mathbf{B}_{k-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{B}_k \end{bmatrix}$$

$$\boldsymbol{w} = \begin{bmatrix} \boldsymbol{w}_{k-1} \\ \hdashline \boldsymbol{w}_k \end{bmatrix}$$

$$\boldsymbol{v} = \begin{bmatrix} \boldsymbol{v}_{k-1} \\ \hdashline \boldsymbol{v}_k \end{bmatrix}$$

A matriz de variâncias-covariâncias das observações será dada por:

$$\mathbf{C}_{\boldsymbol{\ell}} = \begin{bmatrix} \mathbf{C}_{\boldsymbol{\ell}_{k-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\ell}_k} \end{bmatrix}.$$

Pretende-se agora minimizar a função

$$\phi = \hat{\boldsymbol{v}}^T \mathbf{P}_{\boldsymbol{\ell}} \hat{\boldsymbol{v}},$$

que explicitamente corresponde a

$$\phi = \begin{bmatrix} \hat{\boldsymbol{v}}_{k-1}^T & \hat{\boldsymbol{v}}_k^T \end{bmatrix} \begin{bmatrix} \mathbf{P}_{\boldsymbol{\ell}_{k-1}} & \mathbf{0} \\ \mathbf{0} & \mathbf{P}_{\boldsymbol{\ell}_k} \end{bmatrix} \begin{bmatrix} \hat{\boldsymbol{v}}_{k-1} \\ \hdashline \hat{\boldsymbol{v}}_k \end{bmatrix} = \hat{\boldsymbol{v}}_{k-1}^T \mathbf{P}_{\boldsymbol{\ell}_{k-1}} \hat{\boldsymbol{v}}_{k-1} + \hat{\boldsymbol{v}}_k^T \mathbf{P}_{\boldsymbol{\ell}_k} \hat{\boldsymbol{v}}_k,$$

sujeita aos constrangimentos

$$\mathbf{A}_{k-1}\hat{\delta} + \mathbf{B}_{k-1}\hat{v}_{k-1} + w_{k-1} = \mathbf{0}$$

$$\mathbf{A}_{k}\hat{\delta} + \mathbf{B}_{k}\hat{v}_{k} + w_{k} = \mathbf{0}.$$

Assim, a função de variação a minimizar é dada por:

$$\phi = \hat{v}_{k-1}^{T}\mathbf{P}_{\ell_{k-1}}\hat{v}_{k-1} + \hat{v}_{k}^{T}\mathbf{P}_{\ell_{k}}\hat{v}_{k} + 2\hat{\mathbf{k}}_{k-1}^{T}\left(\mathbf{A}_{k-1}\hat{\delta} + \mathbf{B}_{k-1}\hat{v}_{k-1} + w_{k-1}\right) + 2\hat{\mathbf{k}}_{k}^{T}\left(\mathbf{A}_{k}\hat{\delta} + \mathbf{B}_{k}\hat{v}_{k} + w_{k}\right)$$

Calculando as diversas derivadas parciais, tem-se:

$$\frac{\partial\phi}{\partial\hat{v}_{k-1}} = 2\hat{v}_{k-1}^{T}\mathbf{P}_{\ell_{k-1}} + 2\hat{\mathbf{k}}_{k-1}^{T}\mathbf{B}_{k-1} = \mathbf{0},$$

ou

$$\boxed{\mathbf{P}_{\ell_{k-1}}\hat{v}_{k-1} + \mathbf{B}_{k-1}^{T}\hat{\mathbf{k}}_{k-1} = \mathbf{0}}. \tag{4.3}$$

$$\frac{\partial\phi}{\partial\hat{v}_{k}} = 2\hat{v}_{k}^{T}\mathbf{P}_{\ell_{k}} + 2\hat{\mathbf{k}}_{k}^{T}\mathbf{B}_{k} = \mathbf{0},$$

ou

$$\boxed{\mathbf{P}_{\ell_{k}}\hat{v}_{k} + \mathbf{B}_{k}^{T}\hat{\mathbf{k}}_{k} = \mathbf{0}}. \tag{4.4}$$

$$\frac{\partial\phi}{\partial\hat{\delta}} = 2\hat{\mathbf{k}}_{k-1}^{T}\mathbf{A}_{k-1} + 2\hat{\mathbf{k}}_{k}^{T}\mathbf{A}_{k} = \mathbf{0},$$

donde

$$\boxed{\mathbf{A}_{k-1}^{T}\hat{\mathbf{k}}_{k-1} + \mathbf{A}_{k}^{T}\hat{\mathbf{k}}_{k} = \mathbf{0}}. \tag{4.5}$$

O sistema de equações a resolver é dado por

$$\left\{\begin{aligned} \mathbf{A}_{k-1}\hat{\delta} + \mathbf{B}_{k-1}\hat{v}_{k-1} + w_{k-1} &= \mathbf{0} \\ \mathbf{A}_{k}\hat{\delta} + \mathbf{B}_{k}\hat{v}_{k} + w_{k} &= \mathbf{0} \\ \mathbf{P}_{\ell_{k-1}}\hat{v}_{k-1} + \mathbf{B}_{k-1}^{T}\hat{\mathbf{k}}_{k-1} &= \mathbf{0} \\ \mathbf{P}_{\ell_{k}}\hat{v}_{k} + \mathbf{B}_{k}^{T}\hat{\mathbf{k}}_{k} &= \mathbf{0} \\ \mathbf{A}_{k-1}^{T}\hat{\mathbf{k}}_{k-1} + \mathbf{A}_{k}^{T}\hat{\mathbf{k}}_{k} &= \mathbf{0} \end{aligned}\right. \tag{4.6}$$

que pode ser escrito matricialmente sob a forma:

$$\left[\begin{array}{c:cccc} \mathbf{P}_{\ell_{k-1}} & \mathbf{0} & \mathbf{B}_{k-1}^{T} & \mathbf{0} & \mathbf{0} \\ \hdashline \mathbf{0} & \mathbf{P}_{\ell_k} & \mathbf{0} & \mathbf{B}_{k}^{T} & \mathbf{0} \\ \mathbf{B}_{k-1} & \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{k-1} \\ \mathbf{0} & \mathbf{B}_{k} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{k} \\ \mathbf{0} & \mathbf{0} & \mathbf{A}_{k-1}^{T} & \mathbf{A}_{k}^{T} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v}_{k-1} \\ \hdashline \hat{v}_{k} \\ \hat{\mathbf{k}}_{k-1} \\ \hat{\mathbf{k}}_{k} \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline \mathbf{0} \\ w_{k-1} \\ w_{k} \\ \mathbf{0} \end{array}\right] = \mathbf{0}. \tag{4.7}$$

A partição estabelecida permite eliminar $\hat{v}_{k-1}$ e o novo sistema de equações será dado por

$$\left[\begin{array}{c:ccc} \mathbf{P}_{\ell_k} & \mathbf{0} & \mathbf{B}_{k}^{T} & \mathbf{0} \\ \hdashline \mathbf{0} & -\mathbf{B}_{k-1}\mathbf{P}_{\ell_{k-1}}^{-1}\mathbf{B}_{k-1}^{T} & \mathbf{0} & \mathbf{A}_{k-1} \\ \mathbf{B}_{k} & \mathbf{0} & \mathbf{0} & \mathbf{A}_{k} \\ \mathbf{0} & \mathbf{A}_{k-1}^{T} & \mathbf{A}_{k}^{T} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{v}_{k} \\ \hdashline \hat{\mathbf{k}}_{k-1} \\ \hat{\mathbf{k}}_{k} \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} \mathbf{0} \\ \hdashline w_{k-1} \\ w_{k} \\ \mathbf{0} \end{array}\right] = \mathbf{0}, \tag{4.8}$$

o que nos permite eliminar $\hat{v}_{k}$. Obtemos:

$$\left[\begin{array}{c:cc} -\mathbf{B}_{k-1}\mathbf{P}_{\ell_{k-1}}^{-1}\mathbf{B}_{k-1}^{T} & \mathbf{0} & \mathbf{A}_{k-1} \\ \hdashline \mathbf{0} & -\mathbf{B}_{k}\mathbf{P}_{\ell_k}^{-1}\mathbf{B}_{k}^{T} & \mathbf{A}_{k} \\ \mathbf{A}_{k-1}^{T} & \mathbf{A}_{k}^{T} & \mathbf{0} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}}_{k-1} \\ \hdashline \hat{\mathbf{k}}_{k} \\ \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w_{k-1} \\ \hdashline w_{k} \\ \mathbf{0} \end{array}\right] = \mathbf{0}. \tag{4.9}$$

Deste sistema de equações eliminamos $\hat{\mathbf{k}}_{k-1}$ e obtemos:

$$\left[\begin{array}{c:c} -\mathbf{B}_{k}\mathbf{P}_{\ell_k}^{-1}\mathbf{B}_{k}^{T} & \mathbf{A}_{k} \\ \hdashline \mathbf{A}_{k}^{T} & \mathbf{A}_{k-1}^{T}\mathbf{M}_{k-1}^{-1}\mathbf{A}_{k-1} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{k}}_{k} \\ \hdashline \hat{\delta} \end{array}\right] + \left[\begin{array}{c} w_{k} \\ \hdashline \mathbf{A}_{k-1}^{T}\mathbf{M}_{k-1}^{-1}w_{k-1} \end{array}\right] = \mathbf{0}, \tag{4.10}$$

sendo $\mathbf{M}_{K-1} = \mathbf{B}_{k-1}\mathbf{P}_{\ell_{k-1}}^{-1}\mathbf{B}_{k-1}^{T}$.

A diferença entre o ajustamento sequencial e o ajustamento com observações adicionais está no passo seguinte. Assim, em vez de eliminar $\hat{\mathbf{k}}_{k}$, como no ajustamento com observações adicionais, eliminamos $\hat{\delta}$ (dado que A_{22} no sistema de equações (4.10) sistema é invertível) e obtemos uma solução para $\hat{\mathbf{k}}_{k}$:

$$\boxed{\hat{\mathbf{k}}_{k} = \left(\mathbf{B}_{k}\mathbf{P}_{\ell_k}^{-1}\mathbf{B}_{k}^{T} + \mathbf{A}_{k}\mathbf{N}_{k-1}^{-1}\mathbf{A}_{k}^{T}\right)^{-1}\left(w_{k} - \mathbf{A}_{k}\mathbf{N}_{k-1}^{-1}\mathbf{u}_{k-1}\right)}, \tag{4.11}$$

onde

$$\mathbf{N}_{k-1} = \mathbf{A}_{k-1}^{T}\mathbf{M}_{k-1}^{-1}\mathbf{A}_{k-1} \qquad (4.12)$$

e

$$\mathbf{u}_{k-1} = \mathbf{A}_{k-1}^{T}\mathbf{M}_{k-1}^{-1}\boldsymbol{w}_{k-1}\,. \qquad (4.13)$$

Da segunda equação do sistema (4.10), vem:

$$\mathbf{N}_{k-1}\hat{\boldsymbol{\delta}} + \mathbf{A}_{k}^{T}\hat{\mathbf{k}}_{k} + \mathbf{u}_{k-1} = \mathbf{0}$$

$$\boxed{\hat{\boldsymbol{\delta}} = -\mathbf{N}_{k-1}^{-1}\mathbf{u}_{k-1} - \mathbf{N}_{k-1}^{-1}\mathbf{A}_{k}^{T}\hat{\mathbf{k}}_{k}}. \qquad (4.14)$$

Mas a expressão $-\mathbf{N}_{k-1}^{-1}\mathbf{u}_{k-1}$ corresponde ao vector de correcções às aproximações iniciais dos parâmetros desconhecidos, relativo à época t_{k-1}, e que vamos designar por $\hat{\boldsymbol{\delta}}_{k-1}$; $\hat{\boldsymbol{\delta}}$, por seu lado, representa o vector de correcções às aproximações iniciais dos parâmetros desconhecidos, relativo à época t_k, ou seja, considerando todas as observações obtidas até esse instante. Designando este último por $\hat{\boldsymbol{\delta}}_{k}$, a Equação (4.14) pode ser escrita sob a forma:

$$\boxed{\hat{\boldsymbol{\delta}}_{k} = \hat{\boldsymbol{\delta}}_{k-1} - \mathbf{N}_{k-1}^{-1}\mathbf{A}_{k}^{T}\hat{\mathbf{k}}_{k}}. \qquad (4.15)$$

Esta expressão permite obter, de modo sequencial, o vector de correcções às estimativas para os parâmetros desconhecidos num dada época t_k, em função do vector de correcções às estimativas para os parâmetros desconhecidos obtido na época anterior, t_{k-1}. Em cada instante, o vector dos parâmetros desconhecidos será dado por:

$$\boxed{\hat{\boldsymbol{x}}_{k} = \hat{\boldsymbol{x}}_{0} + \hat{\boldsymbol{\delta}}_{k}}. \qquad (4.16)$$

Os resíduos das observações relativos à época t_k obtêm-se da primeira equação do sistema (4.8):

$$\boxed{\hat{\boldsymbol{v}}_{k} = -\mathbf{P}_{\ell_k}^{-1}\mathbf{B}_{k}^{T}\hat{\mathbf{k}}_{k}}. \qquad (4.17)$$

Os resíduos das observações relativos à época t_{k-1} obtêm-se da primeira equação do sistema (4.7):

$$\boxed{\hat{\boldsymbol{v}}_{k-1} = -\mathbf{P}_{\ell_{k-1}}^{-1}\mathbf{B}_{k-1}^{T}\hat{\mathbf{k}}_{k-1}}. \tag{4.18}$$

Os dois vectores de observações ajustadas são:

$$\boxed{\hat{\boldsymbol{\ell}}_{k-1} = \boldsymbol{\ell}_{k-1} + \hat{\boldsymbol{v}}_{k-1}} \tag{4.19}$$

$$\boxed{\hat{\boldsymbol{\ell}}_{k} = \boldsymbol{\ell}_{k} + \hat{\boldsymbol{v}}_{k}}. \tag{4.20}$$

Matriz de variâncias-covariâncias dos parâmetros ajustados

Necessitamos agora de determinar uma expressão sequencial para a matriz de variâncias-covariâncias dos parâmetros ajustados. Para tal, usando a Equação (4.11), começamos por reescrever a Equação (4.15) sob a forma

$$\hat{\boldsymbol{\delta}}_k = \hat{\boldsymbol{\delta}}_{k-1} - \mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\left(\mathbf{M}_k + \mathbf{A}_k\mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\right)^{-1}\left(\boldsymbol{w}_k + \mathbf{A}_k\hat{\boldsymbol{\delta}}_{k-1}\right),$$

ou

$$\hat{\boldsymbol{\delta}}_k = \left[\mathbf{I} - \mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\left(\mathbf{M}_k + \mathbf{A}_k\mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\right)^{-1}\mathbf{A}_k\right]\hat{\boldsymbol{\delta}}_{k-1} - \mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\left(\mathbf{M}_k + \mathbf{A}_k\mathbf{N}_{k-1}^{-1}\mathbf{A}_k^T\right)^{-1}\boldsymbol{w}_k, \tag{4.21}$$

com $\mathbf{M}_K = \mathbf{B}_k\mathbf{P}_{\ell_k}^{-1}\mathbf{B}_k^T$.

Dado que $\hat{\boldsymbol{\delta}}_{k-1}$ é idêntico a $\hat{\boldsymbol{\delta}}$ num ajustamento combinado, a matriz de variâncias-covariâncias para $\hat{\boldsymbol{\delta}}_{k-1}$ terá a mesma expressão - Equação (2.71) -, pelo que podemos considerar que

$$\mathbf{C}_{\hat{\delta}_k} = \mathbf{J}\mathbf{C}_{\ell_k}\mathbf{J}^T = \left[\begin{array}{c:c} \dfrac{\partial\hat{\boldsymbol{\delta}}_k}{\partial\hat{\boldsymbol{\delta}}_{k-1}} & \dfrac{\partial\hat{\boldsymbol{\delta}}_k}{\partial\boldsymbol{\ell}_k}\end{array}\right]\left[\begin{array}{c:c} \sigma_0^2\mathbf{N}_{k-1}^{-1} & 0 \\ \hdashline 0 & \sigma_0^2\mathbf{P}_{\ell_k}^{-1}\end{array}\right]\left[\begin{array}{c} \left(\dfrac{\partial\hat{\boldsymbol{\delta}}_k}{\partial\hat{\boldsymbol{\delta}}_{k-1}}\right)^T \\ \hdashline \left(\dfrac{\partial\hat{\boldsymbol{\delta}}_k}{\partial\boldsymbol{\ell}_k}\right)^T\end{array}\right],$$

ou seja,

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_k} = \sigma_0^2 \left(\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \hat{\boldsymbol{\delta}}_{k-1}} \right) \mathbf{N}_{k-1}^{-1} \left(\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \hat{\boldsymbol{\delta}}_{k-1}} \right)^T + \sigma_0^2 \left(\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \boldsymbol{\ell}_k} \right) \mathbf{P}_{\ell_k}^{-1} \left(\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \boldsymbol{\ell}_k} \right)^T .$$

Usando a seguinte matriz auxiliar

$$\mathbf{Z} = \left(\mathbf{M}_k + \mathbf{A}_k \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \right),$$

tem-se

$$\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \hat{\boldsymbol{\delta}}_{k-1}} = \mathbf{I} - \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{A}_k ,$$

e

$$\frac{\partial \hat{\boldsymbol{\delta}}_k}{\partial \boldsymbol{\ell}_k} = -\mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{B}_K ,$$

pelo que

$$\mathbf{C}_{\hat{\boldsymbol{\delta}}_k} = \sigma_0^2 \left(\mathbf{I} - \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{A}_k \right) \mathbf{N}_{k-1}^{-1} \left(\mathbf{I} - \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{A}_k \mathbf{N}_{k-1}^{-1} \right) + \sigma_0^2 \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{M}_K \mathbf{Z}^{-1} \mathbf{A}_k \mathbf{N}_{k-1}^{-1} .$$

Desenvolvendo e simplificando, tem-se:

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}_k} = \sigma_0^2 \mathbf{N}_{k-1}^{-1} - \sigma_0^2 \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \mathbf{Z}^{-1} \mathbf{A}_k \mathbf{N}_{k-1}^{-1}}, \tag{4.22}$$

ou seja,

$$\boxed{\mathbf{C}_{\hat{\boldsymbol{\delta}}_k} = \mathbf{C}_{\hat{\boldsymbol{\delta}}_{k-1}} - \sigma_0^2 \left(\mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \left(\mathbf{B}_k \mathbf{P}_{\ell_k}^{-1} \mathbf{B}_k^T + \mathbf{A}_k \mathbf{N}_{k-1}^{-1} \mathbf{A}_k^T \right)^{-1} \mathbf{A}_k \mathbf{N}_{k-1}^{-1} \right)}. \tag{4.23}$$

Obtivemos, assim, a matriz de variâncias-covariâncias para as correcções aos parâmetros, à época t_k, em função da matriz de variâncias-covariâncias obtida na época anterior, como se pretendia.

Como é habitual, a matriz de variâncias-covariâncias dos parâmetros será:

$$\boxed{\mathbf{C}_{\hat{x}_k} = \mathbf{C}_{\hat{\boldsymbol{\delta}}_k}} \tag{4.24}$$

5. INTRODUÇÃO AO FILTRO DE KALMAN

O filtro de Kalman foi desenvolvido por R. E. Kalman em 1960. O filtro de Kalman é um método de estimação estocástica que proporciona um meio (recursivo) computacional eficiente e versátil para estimar um estado de um processo, minimizando a soma quadrática dos resíduos.

O filtro de Kalman combina as saídas de um **sensor** (que tem associado uma incerteza de medição, designada na aplicação de filtros como "ruído" associado) de forma a estimar o (**vector de**) **estado** de um sistema com um dado **factor de incerteza dinâmica**.

- <u>Exemplos de sensores</u>: receptores GPS, sistemas de navegação inercial, acelerómetros, relógios, altímetros, ajudas à navegação.
- <u>Exemplos de estado do sistema</u>: posição, velocidade, aceleração, atitude.
  O vector de estado também pode conter parâmetros adicionais como, por exemplo, atrasos atmosféricos (troposfera, ionosfera) associados à propagação de sinais electromagnéticos utilizados por técnicas da Geodesia Espacial, e factores de escala ou enviesamentos (*bias*) em acelerómetros e giroscópios.
- <u>Exemplos de factores de incerteza dinâmica</u>: perturbações não previstas do veículo hospedeiro, causadas pelo próprio utilizador ou pelo meio ambiente (vento, correntes, alterações da topografia do terreno, por exemplo), bem como variações imprevisíveis dos parâmetros do sensor.

O filtro de Kalman é utilizado em diferentes áreas da Ciência, tais como:

- Análise e controlo de sistemas
- Navegação e controlo de estabilidade aérea
- Sistemas inerciais
- Veículos de navegação autónomos
- Controlo de processos químicos
- Processamento e análise de sinais e imagens

Funcionamento do Filtro de Kalman

O filtro de Kalman pretende resolver o problema de estimar o (vector de) estado $\boldsymbol{x} \in \Re^{n}$ de um processo controlado (por exemplo, sujeito a uma aceleração que força o sistema), discreto e linear, sujeito à Equação (5.1),

$$\boxed{\boldsymbol{x}_k = \mathbf{A}\boldsymbol{x}_{k-1} + \mathbf{B}\boldsymbol{u}_{k-1} + \boldsymbol{w}_{k-1}}, \tag{5.1}$$

que constitui o **modelo do sistema**, a partir de um vector de observações (medições) $\mathbf{z} \in \Re^{m}$, relacionado com o vector de estado através da Equação (5.2):

$$\boxed{\mathbf{z}_k = \mathbf{H}x_k + \boldsymbol{v}_k}, \tag{5.2}$$

que constitui o **modelo de observação**, onde as novas variáveis introduzidas são:

$\boldsymbol{w} \cap \mathrm{N}(0, \mathbf{Q})$ **ruído do processo**, com valor médio zero e variância constante, Q. Este tipo de ruído é também designado por **ruído gaussiano** (distribuição normal) ou **ruído branco**

$\boldsymbol{v} \cap \mathrm{N}(0, \mathbf{R})$ **ruído das observações**, com valor médio zero e variância constante, **R**, e <u>independente</u> de ***w***

$\mathbf{A}_{n \times n}$ matriz que relaciona o estado num passo k-1 com o passo k, na ausência de uma função de controlo ou ruído de processo (pode variar de passo para passo, mas assume-se constante)

$\boldsymbol{u} \in \Re^{l}$ vector de funções de controlo (opcional)

$\mathbf{B}_{n \times l}$ Matriz (opcional) que relaciona o controlo de entrada ***u*** com o estado ***x***

$\mathbf{H}_{m \times n}$ Matriz de observação ou matriz de sensibilidade das medições, que relaciona o estado com as observações (pode variar de passo para passo, mas assume-se constante).

Embora seja fácil reconhecer a equivalência de algumas destas variáveis com as utilizadas anteriormente (por exemplo, **H**=**A**, **R**=$\mathbf{C}_{\ell}$ e z=ℓ), optou-se por manter uma notação para esta secção mais próxima da literatura sobre o filtro de Kalman, de modo a facilitar a leitura, embora esta notação seja também variável.

O filtro de Kalman mantém dois tipos de variáveis:

1. Estimativa $\hat{\boldsymbol{x}}$ do vector de estado $\boldsymbol{x}$

   As componentes do vector de estado estimado podem incluir as variáveis de interesse (e.g. posição e velocidade) e parâmetros inúteis (e.g. parâmetros troposféricos).

2. Estimativa da incerteza associada a $\hat{\boldsymbol{x}}$, $\mathbf{P}$

   A incerteza é modelada através da matriz de variâncias-covariâncias dos parâmetros, que tem como definição:

   $$\mathbf{P} \overset{\text{def}}{=} \mathrm{E}\left[\left(\hat{\boldsymbol{x}}-\boldsymbol{x}\right)\left(\hat{\boldsymbol{x}}-\boldsymbol{x}\right)^{\mathrm{T}}\right],$$

   onde E é o operador esperança matemática e $\left(\hat{\boldsymbol{x}}-\boldsymbol{x}\right)$ é a estimativa do erro.

A combinação dos dados novos provenientes do sensor com uma estimativa *a priori*, de forma a obter uma nova estimativa (*a posteriori*), é feita recorrendo a uma matriz de pesos optimizada, denominada **ganho de Kalman**.

O filtro de Kalman é um processo em dois passos, um de **previsão** (**actualização temporal**) e outro de **correcção** (**actualização observacional**, **de medição** ou ***a posteriori***).

Durante a correcção, o filtro actualiza a estimativa do estado e incerteza associada, com base em nova informação proveniente dos sensores (ou seja, utilizando as observações).

Durante a previsão, o filtro actualiza a estimativa do estado e incerteza associada, com base nos efeitos das incertezas dinâmicas do sistema para o período compreendido entre as medições.

O funcionamento do filtro assenta numa série de pressupostos:

- o sistema de equações deve ser linear ou, pelo menos, linearizável em torno de valores aproximados (tal como acontece num ajustamento por mínimos quadrados).
- o ruído que afecta o sistema e as medições é gaussiano (média zero e variância constante);
- os ruídos do sistema e das medições não são correlacionados.

As principais características deste método de estimação são as seguintes:

- algoritmo recursivo de processamento de dados;
- estimador linear óptimo, que opera segundo o critério de minimização da forma quadrática dos resíduos das observações;
- incorpora toda a informação que lhe seja facultada, sob a forma de:
    - modelos da dinâmica linear do sistema e dos dispositivos de medição;
    - descrição estatística dos ruídos, erros de medição e incertezas nos modelos;
    - quaisquer informações disponíveis sobre condições iniciais das variáveis de interesse.

Equações do Filtro

As equações do filtro de Kalman são as seguintes:

1) *Fase de previsão*

a. Previsão do vector de estado

$$\boxed{\hat{\boldsymbol{x}}_k(-) = \mathbf{A}\hat{\boldsymbol{x}}_{k-1}(+) + \mathbf{B}\boldsymbol{u}_{k-1}(+)} \tag{5.3}$$

b. Previsão da matriz de variâncias-covariâncias

$$\boxed{\mathbf{P}_{\mathrm{k}}(-) = \mathbf{A}\mathbf{P}_{\mathrm{k-1}}(+)\mathbf{A}^{\mathrm{T}} + \mathbf{Q}} \tag{5.4}$$

2) *Fase de correcção (actualização das medições ou observações)*

a. Cálculo do ganho de Kalman

$$\boxed{\mathbf{K}_k = \mathbf{P}_k(-)\mathbf{H}^T \left(\mathbf{H}\ \mathbf{P}_k(-)\mathbf{H}^T + \mathbf{R}\ \right)^{-1}} \qquad (5.5)$$

b. Estimativa do estado corrigida:

$$\boxed{\hat{\boldsymbol{x}}_k(+) = \hat{\boldsymbol{x}}_k(-) + \mathbf{K}_k\left(\boldsymbol{z}_k - \mathbf{H}\ \hat{\boldsymbol{x}}_k(-)\right)} \qquad (5.6)$$

c. Matriz de variâncias-covariâncias corrigida:

$$\boxed{\mathbf{P}_k(+) = \mathbf{P}_k(-) - \mathbf{K}_k\mathbf{H}\mathbf{P}_k(-)} \qquad (5.7)$$

onde

$\mathbf{H}\hat{\boldsymbol{x}}(-)$ medição de previsão

$z_k - \mathbf{H}\hat{\boldsymbol{x}}(-)$ vector de inovações (resíduos), correspondente à diferença entre o vector de medição ($\boldsymbol{z}_k$) e a medição de previsão

$\mathbf{K}_K$ ganho de Kalman (ganho ou factor de mistura)

$\mathbf{P}_k(-)$ matriz de variâncias-covariâncias da estimação *a priori* ou de previsão (-)

$\mathbf{P}_k(+)$ matriz de variâncias-covariâncias da estimação *a posteriori* ou corrigida (+)

$\mathbf{Q}$ matriz de variâncias-covariâncias do ruído perturbador dinâmico

$\mathbf{R}$ matriz de variâncias-covariâncias do ruído do sensor ou incertezas das medições

$\hat{\boldsymbol{x}}(-)$ estimativa *a priori* ou de previsão para o vector de estado

$\hat{\boldsymbol{x}}(+)$ estimativa *a posteriori* ou de correcção para o vector de estado

As principais fases da implementação do filtro são:

- Definição do vector de estado $\boldsymbol{x}$.
- Identificação da matriz **A** (dinâmica de transição de estado).

- Identificação da matriz **H** (relação entre as observações e os parâmetros do estado do sistema).
- Avaliação das incertezas no modelo do sistema e dos instrumentos de medição e construção das matrizes de variâncias-covariâncias.
- Estabelecimento das condições iniciais.