

EE210A: Adaptation and Learning

Professor Ali H. Sayed

LECTURE #19

ORDER-RECURSIVE LEAST-SQUARES

Sections in order: 40.1-40.5

MOTIVATION

The recursive least-squares algorithms described so far in Parts VIII (*Least-Squares Methods*) and IX (*Fast RLS Algorithms*), including array variants and fast least-squares variants, are usually qualified as *fixed-order* algorithms. The qualification "fixed-order" means that, from one iteration to another, these implementations propagate quantities that relate to estimation problems of fixed-order.

In this part, we shall study RLS algorithms that are *order-recursive* in nature, as opposed to fixed-order. They are widely known as lattice filters and have several desirable properties such as improved numerical behavior, stability, modularity, in addition to computational efficiency. In these implementations, least-squares problems of increasing orders are solved successively so that, in addition to time-updates, the lattice filters rely heavily on order-updates for various quantities.

NOTATION

Notation for Order-Recursive Problems

To study order-recursive problems, it is necessary to adjust the notation in order to be able to indicate *both* the size of a variable and the time instant at which it becomes available. For example, when referring to a weight vector w_i at time i, we shall write $w_{M,i}$ with two subscripts, M and i. The first subscript, M, is used to indicate that the weight vector is of size M or, equivalently, that it is computed as the solution to a least-squares problem of order M, as in (40.1) and (40.3) below. The second subscript, i, is used to indicate that the weight vector is dependent on data up to time i and, therefore, becomes available at time i.

In a similar vein, we shall write $H_{M,i}$ instead of H_i to refer to a data matrix with column dimension M and with data up to time i. Similarly, we shall write Π_M instead of Π to refer to an $M \times M$ regularization matrix. With this notation, we can now provide a brief review of the regularized least-squares problem.

40.1 MOTIVATION FOR LATTICE FILTERS

So consider a collection of $(i+1)$ data $\{d(j), u_{M,j}\}_{j=0}^{i}$ and introduce the observation vector y_i and the data matrix $H_{M,i}$ defined by

$$y_i = \begin{bmatrix} d(0) \\ d(1) \\ \vdots \\ d(i) \end{bmatrix}, \qquad H_{M,i} = \begin{bmatrix} u_{M,0} \\ u_{M,1} \\ \vdots \\ u_{M,i} \end{bmatrix}$$

The exponentially-weighted least-squares problem of order M seeks the $M \times 1$ column vector w that solves (cf. Sec. 30.6):

$$\min_{w_M} \left[\lambda^{i+1} w_M^* \Pi_M w_M \; + \; (y_i - H_{M,i} w_M)^* \Lambda_i (y_i - H_{M,i} w_M)\right] \qquad (40.1)$$

where Π_M is an $M \times M$ positive-definite regularization matrix. In the sequel, we shall choose Π_M in a manner similar to the fast array method of Chapter 37 (cf. (37.13)), namely,

$$\boxed{\Pi_M \; = \; \eta^{-1} \text{diag}\{\lambda^{-2}, \lambda^{-3}, \ldots, \lambda^{-(M+1)}\}} \qquad (40.2)$$

MOTIVATION

Moreover,

$$\Lambda_i = \text{diag}\{\lambda^i, \lambda^{i-1}, \ldots, \lambda, 1\}$$

is a diagonal weighting matrix, defined in terms of a forgetting factor λ that satisfies $0 \ll \lambda \leq 1$. It is sometimes convenient to rewrite (40.1) more explicitly in terms of the individual data $\{d(j), u_{M,j}\}$ as follows:

$$\min_{w_M} \left[\lambda^{i+1} w_M^* \Pi_M w_M \; + \; \sum_{j=0}^{i} \lambda^{i-j} |d(j) - u_{M,j} w_M|^2 \right] \tag{40.3}$$

We denote the solution of (40.3) by $w_{M,i}$ and we already know that it is given by (cf. Thm. 29.5):

$$\boxed{w_{M,i} \; = \; P_{M,i} H_{M,i}^* \Lambda_i y_i} \tag{40.4}$$

where

$$\boxed{P_{M,i} = (\lambda^{i+1} \Pi_M + H_{M,i}^* \Lambda_i H_{M,i})^{-1}} \tag{40.5}$$

MOTIVATION

The regularization term $\lambda^{i+1}\Pi_M$ guarantees an invertible coefficient matrix, i.e., an invertible $P_{M,i}$. In the absence of regularization (i.e., when $\Pi_M = 0$), we would need to assume that $H_{M,i}$ has full-column rank so that $H^*_{M,i}\Lambda_i H_{M,i}$ is invertible. Observe that the regularization matrix in (40.5) has the form

$$\lambda^{i+1}\Pi_M = \eta^{-1}\text{diag}\{\lambda^{i-1}, \lambda^{i-2}, \ldots, \lambda^{i-M}\}$$

We further let $\widehat{y}_{M,i}$ denote the estimate of y_i,

$$\boxed{\widehat{y}_{M,i} = H_{M,i}w_{M,i}} \tag{40.6}$$

and we refer to $\widehat{y}_{M,i}$ as the regularized projection (or simply projection) of y_i onto the range space of $H_{M,i}$, written as $\mathcal{R}(H_{M,i})$. Recall from Sec. 29.5 that, when $H_{M,i}$ has full-column rank, the projection matrix onto $\mathcal{R}(H_{M,i})$ is $\mathcal{P}_H = H_{M,i}(H^*_{M,i}H_{M,i})^{-1}H^*_{M,i}$. For the regularized problem (40.1), we have $\widehat{y}_{M,i} = H_{M,i}P_{M,i}H^*_{M,i}\Lambda_i y_i$. Although the matrix $H_{M,i}P_{M,i}H^*_{M,i}$ is not an actual projection matrix, we shall still refer to $\widehat{y}_{M,i}$ as the (regularized) projection of y_i onto $\mathcal{R}(H_{M,i})$ for ease of reference.

MOTIVATION

We also define two error vectors: the *a posteriori* and *a priori* error vectors:

$$\boxed{r_{M,i} = y_i - H_{M,i}w_{M,i}, \qquad e_{M,i} = y_i - H_{M,i}w_{M,i-1}} \tag{40.7}$$

where $w_{M,i-1}$ is the solution to a least-squares problem similar to (40.1) and (40.3) with data up to time $i-1$ and with λ^{i+1} replaced by λ^i, i.e.,

$$\min_{w_M} \left[\lambda^i w_M^* \Pi_M w_M \; + \; \sum_{j=0}^{i-1} \lambda^{i-1-j} |d(j) - u_{M,j}w_M|^2 \right] \implies w_{M,i-1}$$

The last entries of the error vectors $\{r_{M,i}, e_{M,i}\}$ at time i are denoted by[21]

$$\left\{ \begin{array}{lcll} r_M(i) & = & d(i) - u_{M,i}w_{M,i} & \text{(}a\ posteriori\text{ error)} \\ e_M(i) & = & d(i) - u_{M,i}w_{M,i-1} & \text{(}a\ priori\text{ error)} \end{array} \right. \tag{40.8}$$

and they are related by the conversion factor, $\gamma_M(i)$,

$$\boxed{r_M(i) = \gamma_M(i)e_M(i)} \tag{40.9}$$

MOTIVATION

which is defined by

$$\boxed{\gamma_M(i) = 1 - u_{M,i}P_{M,i}u_{M,i}^* \;=\; \frac{1}{1+\lambda^{-1}u_{M,i}P_{M,i-1}u_{M,i}^*}} \tag{40.10}$$

Moreover, the minimum cost of the least-squares problem (40.1) is given by (cf. Thm. 29.5):

$$\boxed{\xi_M(i) = y_i^*\Lambda_i r_{M,i} \;=\; y_i^*\Lambda_i[y_i - H_{M,i}w_{M,i}]} \tag{40.11}$$

We also know from Alg. 30.2 that RLS allows us to update $w_{M,i}$ and $\xi_M(i)$ recursively as follows:

MOTIVATION

We also know from Alg. 30.2 that RLS allows us to update $w_{M,i}$ and $\xi_M(i)$ recursively as follows:

$$\left\{\begin{array}{lcl} \gamma_M^{-1}(i) & = & 1+\lambda^{-1}u_{M,i}P_{M,i-1}u_{M,i}^* \\ g_{M,i} & = & \lambda^{-1}\gamma_M(i)P_{M,i-1}u_{M,i}^* \\ e_M(i) & = & d(i)-u_{M,i}w_{M,i-1} \\ w_{M,i} & = & w_{M,i-1}+g_{M,i}e_M(i) \\ P_{M,i} & = & \lambda^{-1}P_{M,i-1}-g_{M,i}g_{M,i}^*/\gamma_M(i) \\ r_M(i) & = & d(i)-u_{M,i}w_{M,i} \\ \xi_M(i) & = & \lambda\xi_M(i-1)\ +\ r_M(i)e_M^*(i) \end{array}\right. \tag{40.12}$$

with initial conditions $w_{M,-1}=0$, $\xi_M(-1)=0$, and $P_{M,-1}=\Pi_M^{-1}$. It also holds that $g_{M,i}=P_{M,i}u_{M,i}^*$.

The RLS algorithm (40.12) allows us to update $w_{M,i-1}$ to $w_{M,i}$, i.e., it only performs a time-update of the weight-vector solution. Here, both $w_{M,i-1}$ and $w_{M,i}$ are $M-$dimensional vectors with the former computed from data up to time $i-1$ while the latter is computed from data up to time i.

MOTIVATION

Now, similar to (40.1), consider a least-squares problem of order $M+1$, i.e.,

$$\min_{w_{M+1}} \left[\lambda^{i+1} w_{M+1}^* \Pi_{M+1} w_{M+1} \;+\; \sum_{j=0}^{i} \lambda^{i-j} |d(j) - u_{M+1,j} w_{M+1}|^2 \right]$$

Its solution is an $(M+1) \times 1$ column vector that we denote by $w_{M+1,i}$. Although an order-update relation that takes $w_{M,i}$ to $w_{M+1,i}$ is possible (recall Lemmas 32.1 and 32.2; see also Prob. X.9), the lattice filters of this chapter are concerned with other kinds of order-update relations.

Specifically, lattice filters are not concerned with the weight vectors themselves, but rather with the corresponding projections $\{\widehat{y}_{M,i}, \widehat{y}_{M+1,i}\}$. So let $d_M(i)$ denote the estimate of $d(i)$ of order M; it is the last entry of $\widehat{y}_{M,i}$, i.e., $d_M(i) = u_{M,i} w_{M,i}$. Likewise, let $d_{M+1}(i)$ denote the estimate of $d(i)$ of order $M+1$, which is the last entry of $\widehat{y}_{M+1,i}$,

$$d_{M+1}(i) = u_{M+1,i} w_{M+1,i} \tag{40.13}$$

MOTIVATION

The corresponding *a posteriori* estimation errors are $r_M(i) = d(i) - d_M(i)$ and $r_{M+1}(i) = d(i) - d_{M+1}(i)$, respectively. It would seem that in order to update $d_M(i)$ to $d_{M+1}(i)$, we may need to order-update $w_{M,i}$ to $w_{M+1,i}$. However, this is not the case. The lattice solutions that we study in this chapter will allow us to update $r_M(i)$ to $r_{M+1}(i)$ *directly* without the need to evaluate the weight vectors $w_{M,i}$ and $w_{M+1,i}$ or even update them. In so doing, the lattice filters will end up being an efficient alternative to RLS; efficient in the sense that their computational cost will be an order of magnitude smaller than that of RLS, namely, $O(M^2)$ vs. $O(M)$ operations per iteration.

40.2 JOINT PROCESS ESTIMATION

We start our derivation of lattice filters by examining the problem of order-updating the projection vector $\widehat{y}_{M,i}$, i.e., of relating $\widehat{y}_{M+1,i}$ to $\widehat{y}_{M,i}$. This problem is known as *joint process estimation*. In order to simplify the presentation, and without loss of generality, we illustrate the arguments and constructions for the case $M = 3$. Later, we show how the results extend to generic M.

Thus assume $M = 3$ and consider the data matrix

$$H_{3,i} = \begin{bmatrix} u(0,0) & u(0,1) & u(0,2) \\ u(1,0) & u(1,1) & u(1,2) \\ u(2,0) & u(2,1) & u(2,2) \\ \vdots & \vdots & \vdots \\ u(i,0) & u(i,1) & u(i,2) \end{bmatrix} = \begin{bmatrix} u_{3,0} \\ u_{3,1} \\ u_{3,2} \\ \vdots \\ u_{3,i} \end{bmatrix} \tag{40.14}$$

The subscript 3 refers to the order of the estimation problem (i.e., to the column dimension of the data matrix), and the subscript i indicates that the data matrix contains data up to time i.

JOINT-PROCESS ESTIMATION

Observe that we are denoting the individual entries of $H_{3,i}$ and, correspondingly, of the regressors $\{u_{3,i}\}$, by $\{u(i,j)\}$ with the first index referring to time and the second index referring to the column position within the regression vector, namely,

$$u_{3,i} = \begin{bmatrix} u(i,0) & u(i,1) & u(i,2) \end{bmatrix}$$

In other words, we are not assuming shift-structure in $u_{3,i}$, i.e., the entries of $u_{3,i}$ are not assumed to be delayed versions of some input sequence. If this were the case, then $H_{3,i}$ would have been of the form

$$H_{3,i} = \begin{bmatrix} u(0) & & \\ u(1) & u(0) & \\ u(2) & u(1) & u(0) \\ \vdots & \vdots & \vdots \\ u(i) & u(i-1) & u(i-2) \end{bmatrix}$$

However, since all results in the sequel, until Sec. 41.1, will hold irrespective of any structure in $u_{3,i}$, we shall proceed with our arguments by treating the general case (40.14).

JOINT-PROCESS ESTIMATION

The (regularized) projection of y_i onto $\mathcal{R}(H_{3,i})$ is given by (cf. (40.6)):

$$\boxed{\widehat{y}_{3,i} = H_{3,i}P_{3,i}H_{3,i}^*\Lambda_i y_i \;=\; H_{3,i}w_{3,i}} \tag{40.15}$$

where

$$\boxed{P_{3,i} = (\lambda^{i+1}\Pi_3 + H_{3,i}^*\Lambda_i H_{3,i})^{-1}, \qquad w_{3,i} = P_{3,i}H_{3,i}^*\Lambda_i y_i} \tag{40.16}$$

and

$$\boxed{\lambda^{i+1}\Pi_3 \;=\; \eta^{-1}\text{diag}\{\lambda^{i-1}, \lambda^{i-2}, \lambda^{i-3}\}} \tag{40.17}$$

We say that $\widehat{y}_{3,i}$ is the third-order projection of y_i onto $\mathcal{R}(H_{3,i})$. Now suppose that one more column is appended to $H_{3,i}$, which then becomes

$$H_{4,i} = \left[\begin{array}{c|c} H_{3,i} & x_{3,i} \end{array}\right] = \left[\begin{array}{ccc|c} u(0,0) & u(0,1) & u(0,2) & u(0,3) \\ u(1,0) & u(1,1) & u(1,2) & u(1,3) \\ u(2,0) & u(2,1) & u(2,2) & u(2,3) \\ \vdots & \vdots & \vdots & \vdots \\ u(i,0) & u(i,1) & u(i,2) & u(i,3) \end{array}\right] \tag{40.18}$$

where we are denoting the last column of $H_{4,i}$ by $x_{3,i}$. The (regularized) projection of the same vector y_i onto the extended range space $\mathcal{R}(H_{4,i})$ is now given by

$$\boxed{\widehat{y}_{4,i} = H_{4,i}P_{4,i}H_{4,i}^*\Lambda_i y_i \;=\; H_{4,i}w_{4,i}} \tag{40.19}$$

where

$$\boxed{P_{4,i} = (\lambda^{i+1}\Pi_4 + H_{4,i}^*\Lambda_i H_{4,i})^{-1}, \qquad w_{4,i} = P_{4,i}H_{4,i}^*\Lambda_i y_i} \tag{40.20}$$

and

$$\boxed{\lambda^{i+1}\Pi_4 \;=\; \eta^{-1}\mathrm{diag}\{\lambda^{i-1}, \lambda^{i-2}, \lambda^{i-3}, \lambda^{i-4}\}} \tag{40.21}$$

Comparing expressions (40.15) and (40.19) for $\{\widehat{y}_{3,i}, \widehat{y}_{4,i}\}$ we see that they differ by virtue of the difference between the data matrices $\{H_{3,i}, H_{4,i}\}$. However, these data matrices are identical except for the last column in $H_{4,i}$. Therefore, it should be possible to relate the projections $\{\widehat{y}_{3,i}, \widehat{y}_{4,i}\}$ and to obtain an order-update relation for them. This process of order-updating the projection of the observation vector is known as *joint process estimation*.

JOINT-PROCESS ESTIMATION

We already studied such order-update problems in Sec. 32.1. Recall that in that section we derived, both algebraically and geometrically, the relations that exist between the (regularized) projection of an observation vector onto a data matrix H and onto its augmented version $[H \quad h]$, for some column h. More specifically, comparing with the statement of Lemma 32.1, we can make the following identifications

$$\left\{\begin{array}{llll} H \longleftarrow H_{3,i} & h \longleftarrow x_{3,i} & P \longleftarrow P_{3,i} & P_z \longleftarrow P_{4,i} \\ \Pi \longleftarrow \lambda^{i+1}\Pi_3 & \sigma \longleftarrow \eta^{-1}\lambda^{i-4} & \gamma \longleftarrow \gamma_3(i) & \gamma_z \longleftarrow \gamma_4(i) \\ \widehat{w} \longleftarrow w_{3,i} & \widehat{w}_z \longleftarrow w_{4,i} & \widetilde{y} \longleftarrow r_{3,i} & \widetilde{y}_z \longleftarrow r_{4,i} \end{array}\right.$$

Therefore, using the result of Lemma 32.1, we can relate the variables of the projection problems that result in $\{\widehat{y}_{3,i}, \widehat{y}_{4,i}\}$ as follows. Let $w^b_{3,i}$ denote the solution of the least-squares problem:

$$\min_{w^b_3} \left[\, \lambda^{i+1} w^{b*}_3 \Pi_3 w^b_3 \; + \; (x_{3,i} - H_{3,i} w^b_3)^* \Lambda_i (x_{3,i} - H_{3,i} w^b_3) \,\right] \qquad (40.22)$$

That is, $w^b_{3,i}$ is the vector that projects $x_{3,i}$ onto $\mathcal{R}(H_{3,i})$,

JOINT-PROCESS ESTIMATION

$$\boxed{w_{3,i}^b = P_{3,i} H_{3,i}^* \Lambda_i x_{3,i}}$$

The subscript 3 refers to an estimation problem of order 3, while the subscript i denotes the use of data up to time i. The superscript b refers to backward projection. The reason for this terminology is that problem (40.22) amounts to estimating the last column of $H_{4,i}$ from its leading columns, $H_{3,i}$. Let $\xi_3^b(i)$ denote the minimum cost of (40.22), i.e.,

$$\boxed{\xi_3^b(i) = x_{3,i}^* \Lambda_i b_{3,i}}$$

where $b_{3,i}$ is the (backward) *a posteriori* error vector that results from projecting $x_{3,i}$ onto $\mathcal{R}(H_{3,i})$,

$$\boxed{b_{3,i} = x_{3,i} - H_{3,i} w_{3,i}^b}$$

We denote the last entry of $b_{3,i}$ by $b_3(i)$ and it refers to the estimation error in estimating the last entry of $x_{3,i}$ from the last row of $H_{3,i}$ (namely, $u_{3,i}$).

JOINT-PROCESS ESTIMATION

Define further the scalar coefficient

$$\boxed{\kappa_3(i) \;\stackrel{\Delta}{=}\; \frac{b_{3,i}^*\Lambda_i y_i}{\eta^{-1}\lambda^{i-4} + \xi_3^b(i)} \;=\; \frac{\rho_3^*(i)}{\eta^{-1}\lambda^{i-4} + \xi_3^b(i)}} \tag{40.23}$$

where

$$\boxed{\rho_3(i) \;\stackrel{\Delta}{=}\; y_i^*\Lambda_i b_{3,i}} \tag{40.24}$$

JOINT-PROCESS ESTIMATION

Then from Lemma 32.1 we conclude that the following order-update relations hold:

$$P_{4,i} = \begin{bmatrix} P_{3,i} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\eta^{-1}\lambda^{i-4} + \xi_3^b(i)} \begin{bmatrix} -w_{3,i}^b \\ 1 \end{bmatrix} \begin{bmatrix} -w_{3,i}^{b^*} & 1 \end{bmatrix} \tag{40.25}$$

$$\widehat{y}_{4,i} = \widehat{y}_{3,i} + \kappa_3(i)\, b_{3,i} \tag{40.26}$$

$$r_{4,i} = r_{3,i} - \kappa_3(i) b_{3,i} \tag{40.27}$$

$$\xi_4(i) = \xi_3(i) - \frac{|\rho_3(i)|^2}{\eta^{-1}\lambda^{i-4} + \xi_3^b(i)} \tag{40.28}$$

$$\gamma_4(i) = \gamma_3(i) - \frac{|b_3(i)|^2}{\eta^{-1}\lambda^{i-4} + \xi_3^b(i)} \tag{40.29}$$

$$w_{4,i} = \begin{bmatrix} w_{3,i} \\ 0 \end{bmatrix} + \kappa_3(i) \begin{bmatrix} -w_{3,i}^b \\ 1 \end{bmatrix} \tag{40.30}$$

We therefore arrived at an order-update relation (40.27) for the *a posteriori* error vectors $\{r_{3,i}, r_{4,i}\}$. It tells us that in order to update $r_{3,i}$ to $r_{4,i}$ we need to know $b_{3,i}$. In the same vein, in order to move forward and update $r_{4,i}$ to $r_{5,i}$ we need $b_{4,i}$ and so on. This means that it is necessary to know how to order-update the backward error vectors as well, which motivates us to examine more closely the backward estimation problem.

40.3 BACKWARD ESTIMATION PROBLEM

For this purpose, we return to the data matrix $H_{3,i}$ in (40.18) and partition it as

$$H_{3,i} = \left[\begin{array}{c|c} x_{0,i} & \bar{H}_{2,i} \end{array}\right]$$

with $x_{0,i}$ denoting its leading column and $\bar{H}_{2,i}$ denoting the remaining columns. In this way, the extended data matrix $H_{4,i}$ of (40.18) can be partitioned as

$$H_{4,i} = \left[\begin{array}{c|c} H_{3,i} & x_{3,i} \end{array}\right] = \left[\begin{array}{ccc} x_{0,i} & \bar{H}_{2,i} & x_{3,i} \end{array}\right] \tag{40.31}$$

with $\{x_{0,i}, x_{3,i}\}$ denoting its leading and trailing columns, and $\bar{H}_{2,i}$ denoting the center columns.

We can then consider two backward estimation problems: one has order 3 and estimates $x_{3,i}$ from $H_{3,i}$, and the other has order 2 and estimates $x_{3,i}$ from $\bar{H}_{2,i}$. The first problem is the one we considered above in (40.22) with regularization matrix $\lambda^{i+1}\Pi_3$ and it leads to the backward residual vector $b_{3,i}$,

$$\boxed{b_{3,i} = x_{3,i} - H_{3,i} w_{3,i}^b} \tag{40.32}$$

BACKWARD ESTIMATION

with the corresponding coefficient matrix

$$\boxed{P_{3,i} = (\lambda^{i+1}\Pi_3 + H_{3,i}^* \Lambda_i H_{3,i})^{-1}} \tag{40.33}$$

The second problem corresponds to solving the following least-squares problem:

$$\min_{w_2^{\bar{b}}} \left[\lambda^i w^{\bar{b}*} \Pi_2 w^{\bar{b}} \; + \; (x_{3,i} - \bar{H}_{2,i} w_2^{\bar{b}})^* \Lambda_i (x_{3,i} - \bar{H}_{2,i} w_2^{\bar{b}}) \right] \tag{40.34}$$

with regularization matrix chosen as

$$\lambda^i \Pi_2 \;=\; \eta^{-1} \text{diag}\{\lambda^{i-2}, \lambda^{i-3}\} \tag{40.35}$$

BACKWARD ESTIMATION

The optimal solution of (40.34) is denoted by $w^{\bar{b}}_{2,i}$ and is given by

$$\boxed{w^{\bar{b}}_{2,i} = \bar{P}_{2,i}\bar{H}^*_{2,i}\Lambda_i x_{3,i}} \tag{40.36}$$

with

$$\boxed{\bar{P}_{2,i} = (\lambda^i\Pi_2 + \bar{H}^*_{2,i}\Lambda_i\bar{H}_{2,i})^{-1}} \tag{40.37}$$

and whose residual vector we denote by

$$\boxed{\bar{b}_{2,i} = x_{3,i} - \bar{H}_{2,i}w^{\bar{b}}_{2,i}}$$

The resulting minimum cost of (40.34) is denoted by $\xi^{\bar{b}}_2(i)$.

BACKWARD ESTIMATION

The resulting minimum cost of (40.34) is denoted by $\xi_2^{\bar{b}}(i)$. The reason for the notation $\bar{b}_{2,i}$ (with an overbar) as opposed to $b_{2,i}$ is that in our development, $b_{2,i}$ would correspond to the residual vector that results from projecting the third column of $H_{3,i}$ onto the range space of its leading two columns. More specifically, denote the columns of $H_{4,i}$ generically by $H_{4,i} = \left[\begin{array}{cccc} m & n & o & p \end{array}\right]$. Then projecting o onto $[m \;\; n]$ results in the residual vector $b_{2,i}$, while projecting p onto $[m \;\; n \;\; o]$ results in the residual vector $b_{3,i}$. Observe that in both cases we start from the initial column m. In contrast, projecting p onto $[n \;\; o]$ results in the residual vector $\bar{b}_{2,i}$. The initial column now is n and, hence, the use of the bar notation to distinguish between both second-order projections: o onto $[m \;\; n]$ and p onto $[n \;\; o]$ — see Fig. 40.1. We shall study more closely later the relation between $\{\bar{b}_{2,i}, b_{2,i}\}$, e.g., in Sec. 41.1 where we show that, when the regressors have shift structure, it will hold that $\bar{b}_{2,i}$ is related to $b_{2,i-1}$. For now, it suffices to proceed with $\bar{b}_{2,i}$.

BACKWARD ESTIMATION

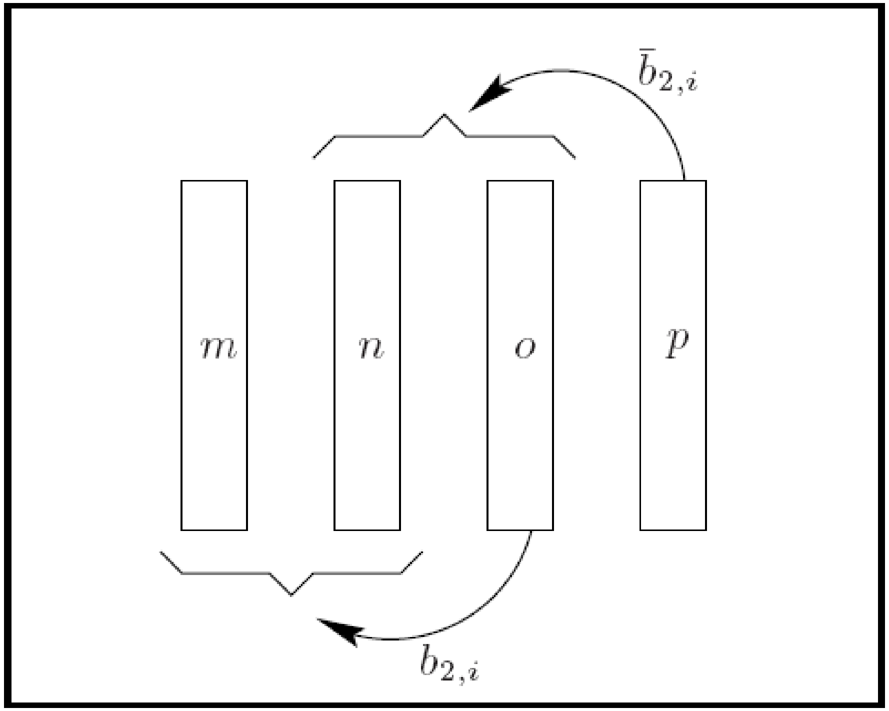


FIGURE 40.1 Two second-order backward projection problems with the corresponding residual vectors.

BACKWARD ESTIMATION

The argument that follows for relating $\bar{b}_{2,i}$ and $b_{3,i}$ is similar to the argument we employed in the previous section for relating $r_{3,i}$ and $r_{4,i}$. Thus note that we are faced with the problem of projecting the same vector $x_{3,i}$ onto the range spaces of two data matrices: one is $\bar{H}_{2,i}$ and the other is $H_{3,i}$, which is obtained from $\bar{H}_{2,i}$ by augmenting it by a column to the left.

We already studied such order-update problems in Sec. 32.2 in some detail. Recall that in that section we derived, both algebraically and geometrically, the relations that exist between the (regularized) projection of an observation vector onto a data matrix H and its augmented version $[h \quad H]$, for some column h. More specifically, comparing with the statement of Lemma 32.2, we can make the following identifications:

$$\left\{ \begin{array}{llll} H \longleftarrow \bar{H}_{2,i} & h \longleftarrow x_{0,i} & P \longleftarrow \bar{P}_{2,i} & P_z \longleftarrow P_{3,i} \\ \Pi \longleftarrow \lambda^i \Pi_2 & \sigma \leftarrow \eta^{-1}\lambda^{i-1} & \gamma \longleftarrow \bar{\gamma}_2(i) & \gamma_z \longleftarrow \gamma_3(i) \\ \widehat{w} \longleftarrow w^{\bar{b}}_{2,i} & \widehat{w}_z \longleftarrow w^b_{3,i} & \widetilde{y} \longleftarrow \bar{b}_{2,i} & \widetilde{y}_z \longleftarrow b_{3,i} \end{array} \right.$$

Therefore, using the result of Lemma 32.2, we can relate the variables of the projection problems that result in $\{\bar{b}_{2,i}, b_{3,i}\}$ as follows.

BACKWARD ESTIMATION

Let $w_{2,i}^f$ denote the solution to the least-squares problem:

$$\min_{w_2^f} \left[\lambda^i w_2^{f*} \Pi_2 w_2^f \; + \; (x_{0,i} - \bar{H}_{2,i} w_2^f)^* \Lambda_i (x_{0,i} - \bar{H}_{2,i} w_2^f) \right] \tag{40.38}$$

which projects the leading column $x_{0,i}$ onto $\mathcal{R}(\bar{H}_{2,i})$, namely,

$$\boxed{w_{2,i}^f = \bar{P}_{2,i} \bar{H}_{2,i}^* \Lambda_i x_{0,i}}$$

The subscript 2 in $w_{2,i}^f$ refers to an estimation problem of order 2, while the subscript i denotes the use of data up to time i. The superscript f refers to forward projection. The reason for this terminology is that the above problem can amounts to estimating the leading column of $H_{3,i}$ from its trailing columns, $\bar{H}_{2,i}$.

BACKWARD ESTIMATION

Let $\xi_2^f(i)$ denote the minimum cost of (40.38), i.e.,

$$\boxed{\xi_2^f(i) = x_{0,i}^* \Lambda_i f_{2,i}}$$

where $f_{2,i}$ is the (forward) *a posteriori* error vector that results from projecting $x_{0,i}$ onto $\mathcal{R}(\bar{H}_{2,i})$,

$$\boxed{f_{2,i} = x_{0,i} - \bar{H}_{2,i} w_{2,i}^f}$$

We denote the last entry of $f_{2,i}$ by $f_2(i)$ and it refers to the estimation error in estimating the last entry of $x_{0,i}$ from the last row of $\bar{H}_{2,i}$.

Define further the scalar coefficient

$$\boxed{\kappa_2^b(i) \;\stackrel{\Delta}{=}\; \frac{f_{2,i}^* \Lambda_i x_{3,i}}{\eta^{-1}\lambda^{i-1} + \xi_2^f(i)} = \frac{\delta_2(i)}{\eta^{-1}\lambda^{i-1} + \xi_2^f(i)}} \tag{40.39}$$

where

$$\boxed{\delta_2(i) \;\stackrel{\Delta}{=}\; f_{2,i}^* \Lambda_i x_{3,i}} \tag{40.40}$$

BACKWARD ESTIMATION

Then from Lemma 32.2 we conclude that the following relations hold:

$$P_{3,i} = \begin{bmatrix} 0 & 0 \\ 0 & \bar{P}_{2,i} \end{bmatrix} + \frac{1}{\eta^{-1}\lambda^{i-1} + \xi_2^f(i)} \begin{bmatrix} 1 \\ -w_{2,i}^f \end{bmatrix} \begin{bmatrix} 1 & -w_{2,i}^{f^*} \end{bmatrix} \tag{40.41}$$

$$b_{3,i} = \bar{b}_{2,i} - \kappa_2^b(i) f_{2,i} \tag{40.42}$$

$$\xi_3^b(i) = \xi_2^{\bar{b}}(i) - \frac{|\delta_2(i)|^2}{\eta^{-1}\lambda^{i-1} + \xi_2^f(i)} \tag{40.43}$$

$$\gamma_3(i) = \bar{\gamma}_2(i) - \frac{|f_2(i)|^2}{\eta^{-1}\lambda^{i-1} + \xi_2^f(i)} \tag{40.44}$$

We therefore arrived at an order-update relation (40.42) for the *a posteriori* backward residual vectors $\{b_{3,i}, \bar{b}_{2,i}\}$. It tells us that in order to update $\bar{b}_{2,i}$ to $b_{3,i}$ we need to know $f_{2,i}$. In the same vein, in order to move forward and update $\bar{b}_{3,i}$ to $b_{4,i}$ we need $f_{3,i}$ and so on. This means that it is necessary to know how to order update the forward error vectors as well, which motivates us to examine more closely the forward estimation problem.

40.4 FORWARD ESTIMATION PROBLEM

To do so, we reconsider the data matrix $H_{4,i}$ in (40.18) and now partition it as

$$H_{4,i} = \left[\begin{array}{c|c} x_{0,i} & \bar{H}_{3,i} \end{array}\right] = \left[\begin{array}{c|cc} x_{0,i} & \bar{H}_{2,i} & x_{3,i} \end{array}\right] \qquad (40.45)$$

where $\bar{H}_{3,i}$ denotes its trailing columns. We then consider two forward estimation problems: one has order 2 and estimates $x_{0,i}$ from $\bar{H}_{2,i}$, and the other has order 3 and estimates $x_{0,i}$ from $\bar{H}_{3,i}$. The first problem is the one we considered above in (40.38) with regularization matrix $\lambda^i\Pi_2$ and leads to the forward residual vector $f_{2,i}$,

$$\boxed{f_{2,i} = x_{0,i} - \bar{H}_{2,i}w^f_{2,i}} \qquad (40.46)$$

with

$$\boxed{\bar{P}_{2,i} = (\lambda^i\Pi_2 + \bar{H}^*_{2,i}\Lambda_i\bar{H}_{2,i})^{-1}} \qquad (40.47)$$

The second problem corresponds to solving the following least-squares problem:

$$\min_{w^f_3} \left[\lambda^i w^{f*}_3\Pi_3 w^f_3 \;+\; (x_{0,i} - \bar{H}_{3,i}w^f_3)^*\Lambda_i(x_{0,i} - \bar{H}_{3,i}w^f_3)\right] \qquad (40.48)$$

FORWARD ESTIMATION

with regularization matrix

$$\lambda^i \Pi_3 \;=\; \eta^{-1}\text{diag}\{\lambda^{i-2}, \lambda^{i-3}, \lambda^{i-4}\} \tag{40.49}$$

The optimal solution of (40.48) is denoted by $w^f_{3,i}$,

$$\boxed{w^f_{3,i} = \bar{P}_{3,i}\bar{H}^*_{3,i}\Lambda_i x_{0,i}} \tag{40.50}$$

with coefficient matrix

$$\boxed{\bar{P}_{3,i} = (\lambda^i\Pi_3 + \bar{H}^*_{3,i}\Lambda_i\bar{H}_{3,i})^{-1}} \tag{40.51}$$

The corresponding residual vector is

$$\boxed{f_{3,i} = x_{0,i} - \bar{H}_{3,i}w^f_{3,i}}$$

and the resulting minimum cost is denoted by $\xi^f_3(i)$.

FORWARD ESTIMATION

Observe that we are now denoting the residual vectors of problems (40.38) and (40.48) by $f_{2,i}$ and $f_{3,i}$, respectively, without the need for the bar notation. Thus note that if we again denote the columns of $H_{4,i}$ generically by $H_{4,i} = \begin{bmatrix} m & n & o & p \end{bmatrix}$. Then projecting m onto $[n \;\; o]$ results in the residual vector $f_{2,i}$, while projecting m onto $[n \;\; o \;\; p]$ results in the residual vector $f_{3,i}$. In both cases, we start from the same initial column n — see Fig. 40.2.

Again, the argument that follows for relating $f_{2,i}$ and $f_{3,i}$ is similar to the arguments we employed in Secs. 40.2 and 40.3 for relating $r_{3,i}$ and $r_{4,i}$, as well as $\bar{b}_{2,i}$ and $b_{3,i}$. Thus note that we are faced with the problem of projecting the same column vector, $x_{0,i}$, onto the range spaces of two data matrices: one is $\bar{H}_{2,i}$ and the other is $\bar{H}_{3,i}$, which is obtained from $\bar{H}_{2.i}$ by augmenting it by a column to the right.

FORWARD ESTIMATION

We studied such order-update problems in Sec. 32.1. Recall that in that section we derived, both algebraically and geometrically, the relations that exist between the (regularized) projection of an observation vector onto a data matrix H and onto its augmented version $[H \quad h]$, for some column h. More specifically, comparing with the statement of Lemma 32.1, we can make the following identifications:

$$\left\{\begin{array}{llllllllllll} H & \longleftarrow & \bar{H}_{2,i} & h & \longleftarrow & x_{3,i} & P & \longleftarrow & \bar{P}_{2,i} & P_z & \longleftarrow & \bar{P}_{3,i} \\ \Pi & \longleftarrow & \lambda^i \Pi_2 & \sigma & \leftarrow & \eta^{-1}\lambda^{i-4} & \gamma & \longleftarrow & \bar{\gamma}_2(i) & \gamma_z & \longleftarrow & \bar{\gamma}_3(i) \\ \widehat{w} & \longleftarrow & w^f_{2,i} & \widehat{w}_z & \longleftarrow & w^f_{3,i} & \widetilde{y} & \longleftarrow & f_{2,i} & \widetilde{y}_z & \longleftarrow & f_{3,i} \end{array}\right.$$

Define further the scalar coefficient

$$\boxed{\kappa^f_2(i) \;\stackrel{\Delta}{=}\; \frac{\bar{b}^*_{2,i}\Lambda_i x_{0,i}}{\eta^{-1}\lambda^{i-4} + \xi^{\bar{b}}_2(i)} = \frac{\delta^*_2(i)}{\eta^{-1}\lambda^{i-4} + \xi^{\bar{b}}_2(i)}} \tag{40.52}$$

FORWARD ESTIMATION

Note that we are using $\delta_2^*(i)$ in the numerator of $\kappa_2^f(i)$, with $\delta_2(i)$ being the coefficient we used in the numerator of $\kappa_2^b(i)$ in (40.39). This is because

$$
\begin{aligned}
\bar{b}_{2,i}^*\Lambda_i x_{0,i} = [x_{3,i} - \bar{H}_{2,i}w_{2,i}^{\bar{b}}]^*\Lambda_i x_{0,i} &= [x_{3,i} - \bar{H}_{2,i}\bar{P}_{2,i}\bar{H}_{2,i}^*\Lambda_i x_{3,i}]^*\Lambda_i x_{0,i} \\
&= x_{3,i}^*\Lambda_i[I - \bar{H}_{2,i}\bar{P}_{2,i}\bar{H}_{2,i}^*\Lambda_i]x_{0,i} \\
&= x_{3,i}^*\Lambda_i[x_{0,i} - \bar{H}_{2,i}w_{2,i}^f] \\
&= x_{3,i}^*\Lambda_i f_{2,i} \qquad (40.53)
\end{aligned}
$$

That is, $\delta_2(i)$ is also given by

$$\boxed{\delta_2(i) = x_{0,i}^*\Lambda_i\bar{b}_{2,i}} \qquad (40.54)$$

FORWARD ESTIMATION

Therefore, using the result of Lemma 32.1, we can relate the variables of the projection problems that result in $\{f_{2,i}, f_{3,i}\}$ as follows:

$$\bar{P}_{3,i} = \begin{bmatrix} \bar{P}_{2,i} & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\eta^{-1}\lambda^{i-4} + \xi_2^{\bar{b}}(i)} \begin{bmatrix} -w_{2,i}^{\bar{b}} \\ 1 \end{bmatrix} \begin{bmatrix} -w_{2,i}^{\bar{b}*} & 1 \end{bmatrix} \tag{40.55}$$

$$f_{3,i} = f_{2,i} - \kappa_2^f(i)\bar{b}_{2,i} \tag{40.56}$$

$$\xi_3^f(i) = \xi_2^f(i) - \frac{|\delta_2(i)|^2}{\eta^{-1}\lambda^{i-4} + \xi_2^{\bar{b}}(i)} \tag{40.57}$$

$$\bar{\gamma}_3(i) = \bar{\gamma}_2(i) - \frac{|\bar{b}_2(i)|^2}{\eta^{-1}\lambda^{i-4} + \xi_2^{\bar{b}}(i)} \tag{40.58}$$

40.5 TIME AND ORDER-UPDATE RELATIONS

We summarize the order-update relations derived so far for the case of a generic order M.

Order-Update of Estimation Errors

Consider several equivalent partitionings of the data matrix $H_{M+1,i}$:

$$\begin{aligned} H_{M+1,i} &= \begin{bmatrix} x_{0,i} & x_{1,i} & \ldots & x_{M,i} \end{bmatrix} \\ &= \begin{bmatrix} H_{M,i} & x_{M,i} \end{bmatrix} = \begin{bmatrix} x_{0,i} & \bar{H}_{M,i} \end{bmatrix} = \begin{bmatrix} x_{0,i} & \bar{H}_{M-1,i} & x_{M,i} \end{bmatrix} \end{aligned}$$

where $\{x_{j,i}\}$ denote the individual columns of $H_{M+1,i}$. Let also $\{u_{M,i}, \bar{u}_{M,i}\}$ denote the last rows of $H_{M,i}$ and $\bar{H}_{M,i}$.

UPDATE RELATIONS

Let further

$$\begin{cases} r_{M,i} &= \textit{a posteriori} \text{ residual from projecting } y_i \text{ onto } H_{M,i} \\ b_{M,i} &= \textit{a posteriori} \text{ residual from projecting } x_{M,i} \text{ onto } H_{M,i} \\ f_{M,i} &= \textit{a posteriori} \text{ residual from projecting } x_{0,i} \text{ onto } \bar{H}_{M,i} \\ \bar{b}_{M,i} &= \textit{a posteriori} \text{ residual from projecting } x_{M+1,i} \text{ onto } \bar{H}_{M,i} \end{cases}$$

where the projection problems for $\{r_{M,i}, b_{M,i}\}$ employ the regularization matrix $\lambda^{i+1}\Pi_M$, while the projection problems for $\{f_{M,i}, \bar{b}_{M,i}\}$ employ the regularization matrix $\lambda^i\Pi_M$. Hence,

$$\begin{cases} r_{M,i} = y_i - H_{M,i}w_{M,i} & b_{M,i} = x_{M,i} - H_{M,i}w^b_{M,i} \\ f_{M,i} = x_{0,i} - \bar{H}_{M,i}w^f_{M,i} & \bar{b}_{M,i} = x_{M+1,i} - \bar{H}_{M,i}w^{\bar{b}}_{M,i} \end{cases} \tag{40.59}$$

where, for example, $w^f_{M,i}$ is the solution to the regularized least-squares problem:

$$\min_{w^f_M} \left[\lambda^i w^{f*}_M \Pi_M w^f_M \;+\; (x_{0,i} - \bar{H}_{M,i}w^f_M)^* \Lambda_i (x_{0,i} - \bar{H}_{M,i}w^f_M) \right] \tag{40.60}$$

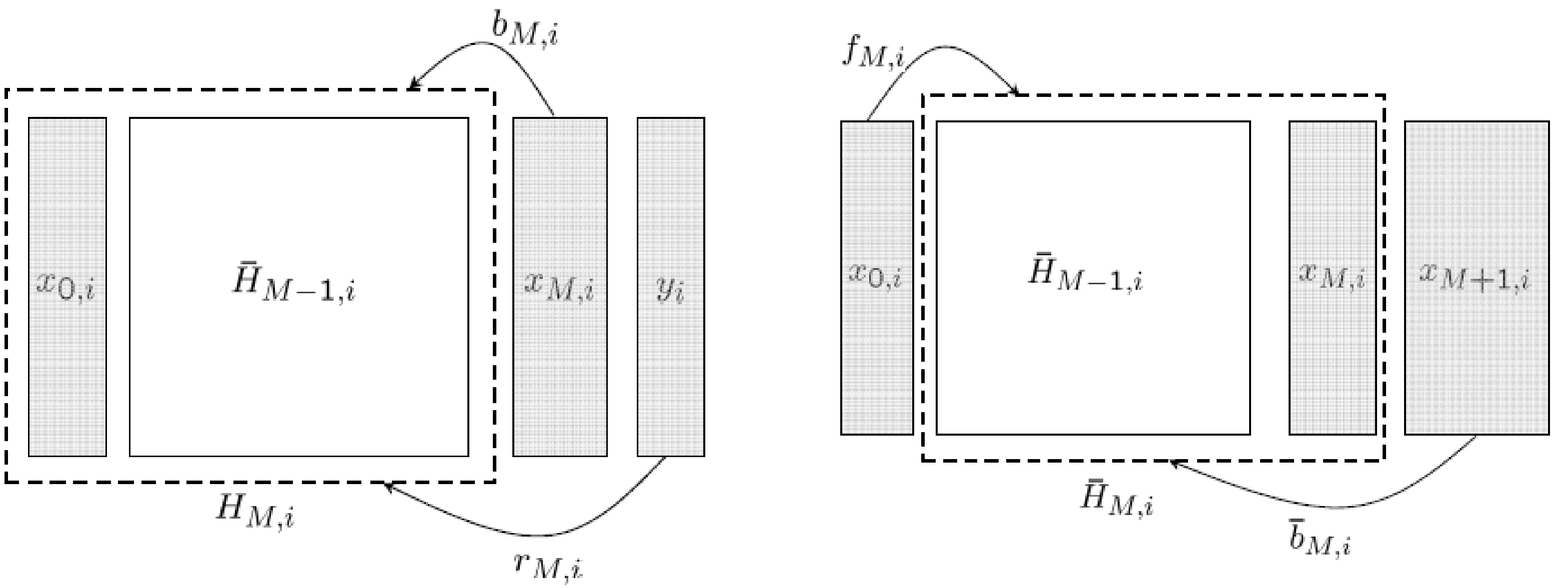
and similarly for $\{w_{M,i}, w^b_{M,i}, w^{\bar{b}}_{M,i}\}$. These constructions are depicted in Fig. 40.3.

UPDATE RELATIONS

These constructions are depicted in Fig. 40.3.

The last entries of $\{r_{M,i}, b_{M,i}, f_{M,i}, \bar{b}_{M,i}\}$ are denoted by

$$\{r_M(i), b_M(i), f_M(i), \bar{b}_M(i)\} \tag{40.61}$$



(a) Backward projection (b) Forward projection

FIGURE 40.3 Projections of $\{x_{0,i}, x_{M,i}, x_{M+1,i}, y_i\}$ onto the relevant data matrices with the resulting residual vectors $\{f_{M,i}, b_{M,i}, \bar{b}_{M,i}, r_{M,i}\}$.

UPDATE RELATIONS

The derivations in the earlier sections show that these residual vectors satisfy updates of the form:

$$r_{M+1,i} = r_{M,i} - \kappa_M(i) b_{M,i}, \quad b_{M+1,i} = \bar{b}_{M,i} - \kappa_M^b(i) f_{M,i}, \quad f_{M+1,i} = f_{M,i} - \kappa_M^f(i) \bar{b}_{M,i} \tag{40.62}$$

where we still need to derive an update for $\bar{b}_{M,i}$. We postpone this discussion to Sec. 41.1 due to its dependence on data structure. From (40.62) we obtain the following relations for the *a posteriori* estimation errors at time i:

$$\left\{ \begin{array}{lcl} r_{M+1}(i) & = & r_M(i) - \kappa_M(i) b_M(i) \\ b_{M+1}(i) & = & \bar{b}_M(i) - \kappa_M^b(i) f_M(i) \\ f_{M+1}(i) & = & f_M(i) - \kappa_M^f(i) \bar{b}_M(i) \end{array} \right. \tag{40.63}$$

where the scaling coefficients $\{\kappa_M(i), \kappa_M^b(i), \kappa_M^f(i)\}$, also called reflection coefficients, are defined as the ratios

$$\left\{ \begin{array}{lcl} \kappa_M(i) & = & \rho_M^*(i)/(\eta^{-1}\lambda^{i-M-1} + \xi_M^b(i)) \\ \kappa_M^b(i) & = & \delta_M(i)/(\eta^{-1}\lambda^{i-1} + \xi_M^f(i)) \\ \kappa_M^f(i) & = & \delta_M^*(i)/(\eta^{-1}\lambda^{i-M-2} + \xi_M^{\bar{b}}(i)) \end{array} \right. \tag{40.64}$$

UPDATE RELATIONS

and the quantities $\{\delta_M(i), \rho_M(i), \xi^b_M(i), \xi^{\bar{b}}_M(i), \xi^f_M(i)\}$ are defined in terms of the inner products

$$\left\{\begin{array}{lllll} \rho_M(i) & = & y_i^*\Lambda_i b_{M,i} & & \\ \delta_M(i) & = & x_{0,i}^*\Lambda_i\bar{b}^*_{M,i} & = & f^*_{M,i}\Lambda_i x_{M+1,i} \\ \xi^f_M(i) & = & x^*_{0,i}\Lambda_i f_{M,i} & & \end{array}\right. \quad \begin{array}{lll} \xi^b_M(i) & = & x^*_{M,i}\Lambda_i b_{M,i} \\ \xi^{\bar{b}}_M(i) & = & x^*_{M+1,i}\Lambda_i\bar{b}_{M,i} \\ \xi_M(i) & = & y_i^*\Lambda_i r_{M,i} \end{array} \qquad (40.65)$$

The quantities $\{\xi_M(i), \xi^b_M(i), \xi^f_M(i), \xi^{\bar{b}}_M(i)\}$ denote the minimum costs of the projection problems that result in $\{r_{M,i}, b_{M,i}, f_{M,i}, \bar{b}_{M,i}\}$. Let further $\{\gamma_M(i), \bar{\gamma}_M(i)\}$ denote the conversion factors associated with the projection problems $\{r_{M,i}, b_{M,i}\}$ and $\{f_{M,i}, \bar{b}_{M,i}\}$ (the first two have the same conversion factor $\gamma_M(i)$, while the last two have the same conversion factor $\bar{\gamma}_M(i)$). That is,

$$\left\{\begin{array}{lll} \gamma_M(i) & = & 1 - u_{M,i}P_{M,i}u^*_{M,i} \\ \bar{\gamma}_M(i) & = & 1 - \bar{u}_{M,i}\bar{P}_{M,i}\bar{u}^*_{M,i} \end{array}\right. \quad \begin{array}{lll} P_{M,i} & = & \left[\lambda^{i+1}\Pi_M + H_{M,i}\Lambda_i H^*_{M,i}\right]^{-1} \\ \bar{P}_{M,i} & = & \left[\lambda^i\Pi_M + \bar{H}_{M,i}\Lambda_i\bar{H}^*_{M,i}\right]^{-1} \end{array}$$

UPDATE RELATIONS

Then the earlier discussions also established the following update relations:

$$\left\{\begin{array}{lcl}
\xi_{M+1}(i) & = & \xi_M(i) - |\rho_M(i)|^2/(\eta^{-1}\lambda^{i-M-1} + \xi^b_M(i)) \\
\xi^b_{M+1}(i) & = & \xi^{\bar{b}}_M(i) - |\delta_M(i)|^2/(\eta^{-1}\lambda^{i-1} + \xi^f_M(i)) \\
\xi^f_{M+1}(i) & = & \xi^f_M(i) - |\delta_M(i)|^2/(\eta^{-1}\lambda^{i-M-2} + \xi^{\bar{b}}_M(i)) \\
 & & \\
\gamma_{M+1}(i) & = & \gamma_M(i) - |b_M(i)|^2/(\eta^{-1}\lambda^{i-M-1} + \xi^b_M(i)) \\
\gamma_{M+1}(i) & = & \bar{\gamma}_M(i) - |f_M(i)|^2/(\eta^{-1}\lambda^{i-1} + \xi^f_M(i)) \\
\bar{\gamma}_{M+1}(i) & = & \bar{\gamma}_M(i) - |\bar{b}_M(i)|^2/(\eta^{-1}\lambda^{i-M-2} + \xi^{\bar{b}}_M(i))
\end{array}\right. \tag{40.66}$$

as well as (cf. (40.30)):[22]

$$w_{M+1,i} = \begin{bmatrix} w_{M,i} \\ 0 \end{bmatrix} + \kappa_M(i)\begin{bmatrix} -w^b_{M,i} \\ 1 \end{bmatrix} \tag{40.67}$$

We still need to show how to update the factors $\{\rho_M(i), \delta_M(i)\}$ in (40.65) in order to arrive at an efficient recursive scheme. The derivation in the next section shows that time-updates for $\{\rho_M(i), \delta_M(i)\}$ are possible regardless of data structure.

UPDATE RELATIONS

Time-Update Relations

Consider first the quantity $\delta_M(i) = x^*_{0,i}\Lambda_i \bar{b}_{M,i}$, which appears in the numerator of $\kappa^f_M(i)$ in (40.64), and introduce the data matrix

$$H_{M+2,i} = \left[\begin{array}{ccc} x_{0,i} & \bar{H}_{M,i} & x_{M+1,i} \end{array} \right]$$

We partition it as

$$H_{M+2,i} = \left[\begin{array}{c|c|c} x_{0,i-1} & \bar{H}_{M,i-1} & x_{M+1,i-1} \\ \hline u(i,0) & \bar{u}_{M,i} & u(i,M+1) \end{array} \right]$$

where we are denoting the last entries of $\{x_{0,i}, x_{M+1,i}\}$ by $\{u(i,0), u(i,M+1)\}$, and the last row of $\bar{H}_{M,i}$ by $\bar{u}_{M,i}$. Consider further

$$\delta_M(i-1) = x^*_{0,i-1}\Lambda_{i-1}\bar{b}_{M,i-1}$$

UPDATE RELATIONS

Now recall that $\bar{b}_{M,i}$ is the residual vector that results from projecting $x_{M+1,i}$ onto $\mathcal{R}(\bar{H}_{M,i})$ with regularization matrix $\lambda^i \Pi_M$. Likewise, $\bar{b}_{M,i-1}$ is the residual vector that results from projecting $x_{M+1,i-1}$ onto $\mathcal{R}(\bar{H}_{M,i-1})$ with regularization matrix $\lambda^{i-1}\Pi_M$. We are therefore faced with the problem of time-updating the inner product $\delta_M(i)$, which is of the same form as the problem studied in Sec. 32.3. More specifically, comparing with the statement of Lemma 32.3 (or with the data matrix (32.48) and its time-updated version (32.50)), we see that we can make the identifications:

$$\left\{\begin{array}{llllll} \bar{H}_{i-1} \longleftarrow \bar{H}_{M,i-1} & & \bar{\gamma}(i) \longleftarrow \bar{\gamma}_M(i) & & \beta(i) \longleftarrow u(i, M+1) \\ x_{i-1} \longleftarrow x_{0,i-1} & & \bar{h}_i \longleftarrow \bar{u}_{M,i} & & \widetilde{\alpha}(i) \longleftarrow f_M(i) \\ \alpha(i) \longleftarrow u(i,0) & & z_{i-1} \longleftarrow x_{M+1,i-1} & & \widetilde{\beta}(i) \longleftarrow b_M(i) \end{array}\right.$$

and arrive at the time-update relation (cf. (32.56)):

$$\boxed{\delta_M(i) = \lambda\delta_M(i-1) + \frac{f_M^*(i)\bar{b}_M(i)}{\bar{\gamma}_M(i)}} \qquad (40.68)$$

UPDATE RELATIONS

Consider now the inner product $\rho_M(i) = y_i^* \Lambda_i b_{M,i}$, which appears in the numerator of $\kappa_M(i)$ in (40.64), and introduce the matrix

$$\begin{bmatrix} y_i & H_{M,i} & x_{M,i} \end{bmatrix}$$

Let us partition it as

$$\left[\begin{array}{c|c|c} y_{i-1} & H_{M,i-1} & x_{M,i-1} \\ \hline d(i) & u_{M,i} & u(i,M) \end{array}\right]$$

Now recall that $b_{M,i}$ is the residual vector that results from projecting $x_{M,i}$ onto $\mathcal{R}(H_{M,i})$ with regularization matrix $\lambda^{i+1}\Pi_M$, while $b_{M,i-1}$ is the residual vector that results from projecting $x_{M,i-1}$ onto $\mathcal{R}(H_{M,i-1})$ with regularization matrix $\lambda^i \Pi_M$. We are therefore faced with the problem of time-updating the inner product $\rho_M(i)$, which is again of the same form as the problem studied earlier in Sec. 32.3. More specifically, comparing with the statement of Lemma 32.3 (or with the data matrix (32.48) and its time-updated version (32.50)), we see that we can make the identifications:

UPDATE RELATIONS

$$\left\{\begin{array}{lll} \bar{H}_{i-1} \longleftarrow H_{M,i-1} & \gamma(i) \longleftarrow \gamma_M(i) & \beta(i) \longleftarrow u(i,M) \\ x_{i-1} \longleftarrow y_{i-1} & \bar{h}_i \longleftarrow u_{M,i} & \widetilde{\alpha}(i) \longleftarrow e_M(i) \\ \alpha(i) \longleftarrow d(i) & z_{i-1} \longleftarrow x_{M,i-1} & \widetilde{\beta}(i) \longleftarrow b_M(i) \end{array}\right.$$

and arrive at the time-update relation (cf. (32.56)):

$$\boxed{\rho_M(i) = \lambda \rho_M(i-1) + \frac{r_M^*(i) b_M(i)}{\gamma_M(i)}} \tag{40.69}$$

We can also obtain time-updates for the minimum costs $\{\xi_M^f(i), \xi_M^b(i), \xi_M^{\bar{b}}(i)\}$ in much the same manner as above. Alternatively, since these variables correspond to the minimum costs of regularized least-squares problems, and since we already know how to time-update such minimum costs (cf. Alg. 30.2), we can readily write

$$\left\{\begin{array}{lll} \xi_M^f(i) & = & \lambda \xi_M^f(i-1) + |f_M(i)|^2 / \bar{\gamma}_M(i) \\ \xi_M^b(i) & = & \lambda \xi_M^b(i-1) + |b_M(i)|^2 / \gamma_M(i) \\ \xi_M^{\bar{b}}(i) & = & \lambda \xi_M^{\bar{b}}(i-1) + |\bar{b}_M(i)|^2 / \bar{\gamma}_M(i) \end{array}\right. \tag{40.70}$$

UPDATE RELATIONS

Table 40.1 collects the various order- and time-update relations derived so far in the chapter. We again emphasize that these relations are independent of any data structure. For convenience of notation, and also in order to save on addition operations, we introduce the *modified* cost variables:

$$\left\{\begin{array}{lcl}\zeta_M^b(i) & \stackrel{\Delta}{=} & \eta^{-1}\lambda^{i-M-1}+\xi_M^b(i)\\ \zeta_M^f(i) & \stackrel{\Delta}{=} & \eta^{-1}\lambda^{i-1}+\xi_M^f(i)\\ \zeta_M^{\bar{b}}(i) & \stackrel{\Delta}{=} & \eta^{-1}\lambda^{i-M-2}+\xi_M^{\bar{b}}(i)\end{array}\right. \tag{40.71}$$

UPDATE RELATIONS

It is easy to verify from the time and order-updates for $\{\xi^b_M(i), \xi^f_M(i), \xi^{\bar{b}}_M(i)\}$ that these modified variables satisfy similar updates, namely

$$\left\{\begin{array}{lcl} \zeta^f_M(i) & = & \lambda\zeta^f_M(i-1) + |f_M(i)|^2/\bar{\gamma}_M(i) \\ \zeta^b_M(i) & = & \lambda\zeta^b_M(i-1) + |b_M(i)|^2/\gamma_M(i) \\ \zeta^{\bar{b}}_M(i) & = & \lambda\zeta^{\bar{b}}_M(i-1) + |\bar{b}_M(i)|^2/\bar{\gamma}_M(i) \\ & & \\ \zeta_{M+1}(i) & = & \zeta_M(i) - |\rho_M(i)|^2/\zeta^b_M(i) \\ \zeta^b_{M+1}(i) & = & \zeta^{\bar{b}}_M(i) - |\delta_M(i)|^2/\zeta^f_M(i) \\ \zeta^f_{M+1}(i) & = & \zeta^f_M(i) - |\delta_M(i)|^2/\zeta^{\bar{b}}_M(i) \end{array}\right. \tag{40.72}$$

albeit with initial conditions

$$\zeta^b_M(-1) = \eta^{-1}\lambda^{-M-2}, \quad \zeta^f_M(-1) = \eta^{-1}\lambda^{-2}, \quad \zeta^{\bar{b}}_M(-1) = \eta^{-1}\lambda^{-M-3} \tag{40.73}$$

while the initial values for the original variables are

$$\xi^b_M(-1) = \xi^f_M(-1) = \xi^{\bar{b}}_M(-1) = 0 \tag{40.74}$$

UPDATE RELATIONS

TABLE 40.1 A listing of the time and order-update relations derived in Secs. 40.2–40.5. All these updates are independent of data structure.

$$
\begin{aligned}
\xi_M^f(i) &= \lambda \xi_M^f(i-1) + |f_M(i)|^2/\bar{\gamma}_M(i) \\
\xi_M^b(i) &= \lambda \xi_M^b(i-1) + |b_M(i)|^2/\gamma_M(i) \\
\xi_M^{\bar{b}}(i) &= \lambda \xi_M^{\bar{b}}(i-1) + |\bar{b}_M(i)|^2/\bar{\gamma}_M(i) \\
\\
\xi_{M+1}(i) &= \xi_M(i) - |\rho_M(i)|^2/\zeta_M^b(i) \\
\xi_{M+1}^b(i) &= \xi_M^{\bar{b}}(i) - |\delta_M(i)|^2/\zeta_M^f(i) \\
\xi_{M+1}^f(i) &= \xi_M^f(i) - |\delta_M(i)|^2/\zeta_M^{\bar{b}}(i) \\
\\
\rho_M(i) &= \lambda \rho_M(i-1) + r_M^*(i) b_M(i)/\gamma_M(i) \\
\delta_M(i) &= \lambda \delta_M(i-1) + f_M^*(i) \bar{b}_M(i)/\bar{\gamma}_M(i) \\
\\
\kappa_M(i) &= \rho_M^*(i)/\zeta_M^b(i) \\
\kappa_M^b(i) &= \delta_M(i)/\zeta_M^f(i) \\
\kappa_M^f(i) &= \delta_M^*(i)/\zeta_M^{\bar{b}}(i) \\
\\
r_{M+1}(i) &= r_M(i) - \kappa_M(i) b_M(i) \\
b_{M+1}(i) &= \bar{b}_M(i) - \kappa_M^b(i) f_M(i) \\
f_{M+1}(i) &= f_M(i) - \kappa_M^f(i) \bar{b}_M(i) \\
\\
\gamma_{M+1}(i) &= \gamma_M(i) - |b_M(i)|^2/\zeta_M^b(i) \\
\gamma_{M+1}(i) &= \bar{\gamma}_M(i) - |f_M(i)|^2/\zeta_M^f(i) \\
\bar{\gamma}_{M+1}(i) &= \bar{\gamma}_M(i) - |\bar{b}_M(i)|^2/\zeta_M^{\bar{b}}(i)
\end{aligned}
$$