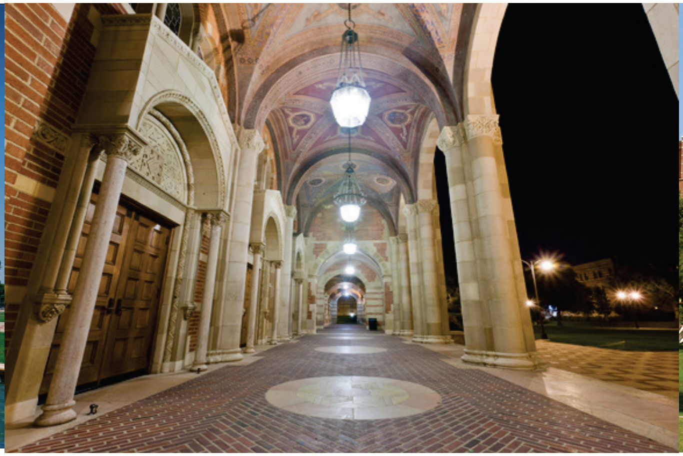

EE210A: Adaptation and Learning

Professor Ali H. Sayed

LECTURE #18

ORDER- AND TIME-UPDATE RELATIONS

Sections in order: 32.1-32.3

32.1 BACKWARD ORDER-UPDATE RELATIONS

Consider a weighted regularized least-squares problem of the form

$$\min_{w} \left[w^*\Pi w + (y - Hw)^* W (y - Hw) \right] \tag{32.1}$$

whose optimal solution is given by (cf. Thm. 29.5):

$$\widehat{w} = (\Pi + H^* W H)^{-1} H^* W y \tag{32.2}$$

Let

$$\boxed{P \stackrel{\Delta}{=} (\Pi + H^* W H)^{-1}} \tag{32.3}$$

In other words, P is the inverse of the coefficient matrix that appears in the normal equations $(\Pi + H^* W H)\widehat{w} = H^* W y$. Then

$$\widehat{w} = P H^* W y$$

BACKWARD UPDATE RELATIONS

and the estimate of y is

$$\widehat{y} = H\widehat{w} \;=\; HPH^*Wy \tag{32.4}$$

We refer to $\widehat{y}$ as the regularized projection (or simply projection) of y onto the range space of H. Recall from Sec. 29.5 that, when H has full column rank, the actual projection matrix onto $\mathcal{R}(H)$ is defined by

$$\mathcal{P}_H = H(H^*H)^{-1}H^*$$

For the regularized problem (32.1), we instead have

$$\widehat{y} = H(\Pi + H^*WH)^{-1}H^*Wy \;=\; HPH^*Wy$$

Although the matrix HPH^*W that multiplies y is not an actual projection matrix (cf. (29.19) and (29.20)), we shall still refer to $\widehat{y}$ as the projection of y onto $\mathcal{R}(y)$ for ease of reference. Note that in both cases, with and without regularization, $\widehat{y} \in \mathcal{R}(H)$.

BACKWARD UPDATE RELATIONS

The resulting residual vector is

$$\boxed{\widetilde{y} = y - H\widehat{w}}$$

and the corresponding minimum cost is (cf. Table 29.3):

$$\boxed{\xi = y^* W \widetilde{y}} \quad (32.5)$$

Introduce the scalar

$$\boxed{\gamma \overset{\Delta}{=} 1 - uPu^*} \quad (32.6)$$

where u denotes the last row of H. This scalar plays an important role in least-squares adaptive filtering, and it is called the conversion factor for reasons to be explained later in Sec. 30.4. We shall not employ γ in any of the arguments below, but will only comment on some of its properties whenever appropriate.

BACKWARD UPDATE RELATIONS

Now assume that we extend the data matrix H in (32.1) by adding one column to it, say h, and consider the extended least-squares problem:

$$\min_{w_z} \left\{ w_z^* \begin{bmatrix} \Pi & 0 \\ 0 & \sigma \end{bmatrix} w_z + \left(y - \begin{bmatrix} H & h \end{bmatrix} w_z\right)^* W \left(y - \begin{bmatrix} H & h \end{bmatrix} w_z\right) \right\} \tag{32.7}$$

We say that the order of the estimation problem has increased by one since we are now estimating y from the column span of the extended data matrix $[H \quad h]$. Of course, the corresponding weight vector, w_z, has one dimension higher than w and, accordingly, we also extend the regularization matrix, Π, by adding a positive scalar σ to it.[16]

The optimal solution of (32.7) is given by (again cf. Thm. 29.5):

$$\widehat{w}_z = \left(\begin{bmatrix} \Pi & 0 \\ 0 & \sigma \end{bmatrix} + \begin{bmatrix} H^* \\ h^* \end{bmatrix} W \begin{bmatrix} H & h \end{bmatrix} \right)^{-1} \begin{bmatrix} H^* \\ h^* \end{bmatrix} W y \tag{32.8}$$

Similarly to P in (32.3), we define

$$\boxed{P_z = \left(\begin{bmatrix} \Pi & 0 \\ 0 & \sigma \end{bmatrix} + \begin{bmatrix} H^* \\ h^* \end{bmatrix} W \begin{bmatrix} H & h \end{bmatrix} \right)^{-1}} \tag{32.9}$$

BACKWARD UPDATE RELATIONS

so that the new estimate of y is given by

$$\widehat{y}_z = \begin{bmatrix} H & h \end{bmatrix} \widehat{w}_z = \begin{bmatrix} H & h \end{bmatrix} P_z \begin{bmatrix} H^* \\ h^* \end{bmatrix} W y \qquad (32.10)$$

We say that $\widehat{y}_z$ is the regularized projection of y onto the range space of the extended data matrix $[H \quad h]$. The resulting residual vector is

$$\widetilde{y}_z = y - \begin{bmatrix} H & h \end{bmatrix} \widehat{w}_z$$

and the corresponding minimum cost is (cf. Table 29.3):

$$\xi_z = y^* W \widetilde{y}_z \qquad (32.11)$$

The associated conversion factor is

$$\gamma_z \overset{\Delta}{=} 1 - \begin{bmatrix} u & \alpha \end{bmatrix} P_z \begin{bmatrix} u^* \\ \alpha^* \end{bmatrix} \qquad (32.12)$$

BACKWARD UPDATE RELATIONS

where $[u \;\; \alpha]$ denotes the last row of $[H \;\; h]$,

$$\left[\begin{array}{cc} H & h \end{array}\right] = \left[\begin{array}{cc} \times & \times \\ \times & \times \\ \hline u & \alpha \end{array}\right]$$

Our objective is to examine the relations between the solutions of the least-squares problems (32.1) and (32.7). We shall use both algebraic and geometric arguments for the sake of illustration. We start with the algebraic argument and later show how the geometry of least-squares theory can be used to arrive at the same conclusions.

Algebraic Argument

The first step toward relating the solution vectors $\{\widehat{w}, \widehat{w}_z\}$ in (32.2) and (32.8) is to relate the coefficient matrices $\{P, P_z\}$ in (32.3) and (32.9). To achieve this, we start by observing from (32.9) that

$$P_z^{-1} = \left[\begin{array}{cc} \Pi & 0 \\ 0 & \sigma \end{array}\right] + \left[\begin{array}{c} H^* \\ h^* \end{array}\right] W \left[\begin{array}{cc} H & h \end{array}\right] = \left[\begin{array}{cc} P^{-1} & H^*Wh \\ h^*WH & \sigma + h^*Wh \end{array}\right] \tag{32.13}$$

BACKWARD UPDATE RELATIONS

where the top leftmost corner entry of P_z^{-1} is seen to be P^{-1},

$$P^{-1} = \Pi + H^*WH$$

In order to relate P_z to P we invoke the easily verifiable matrix identity:

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} -A^{-1}B \\ \text{I} \end{bmatrix} (D - CA^{-1}B)^{-1} \begin{bmatrix} -CA^{-1} & \text{I} \end{bmatrix} \tag{32.14}$$

which relates the inverse of a block matrix to the inverse of its top leftmost corner block. Applying this identity to (32.13) we obtain

$$P_z = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\nu} \begin{bmatrix} -PH^*Wh \\ 1 \end{bmatrix} \begin{bmatrix} -h^*WHP & 1 \end{bmatrix} \tag{32.15}$$

where the scalar ν is given by

$$\nu \;=\; \sigma + h^*Wh - h^*WHPH^*Wh = \sigma + h^*W[h - HPH^*Wh] \tag{32.16}$$

BACKWARD UPDATE RELATIONS

The quantity PH^*Wh that appears in (32.15) and (32.16) can be interpreted as the weight vector that solves the following regularized least-squares problem:

$$\min_{w^b} \left[w^{b*}\Pi w^b + (h - Hw^b)^* W (h - Hw^b) \right] \tag{32.17}$$

Indeed, the solution of (32.17) is $\widehat{w}^b = PH^*Wh$, and it corresponds to the weight vector that projects (in a regularized manner) the column vector h onto the column span of H. This projection problem is the reason for the title of this subsection, namely, *backward projection*. The terminology refers to the fact that we are projecting h onto the preceding columns of $[H \quad h]$, as indicated schematically in Fig. 32.1.

BACKWARD UPDATE RELATIONS

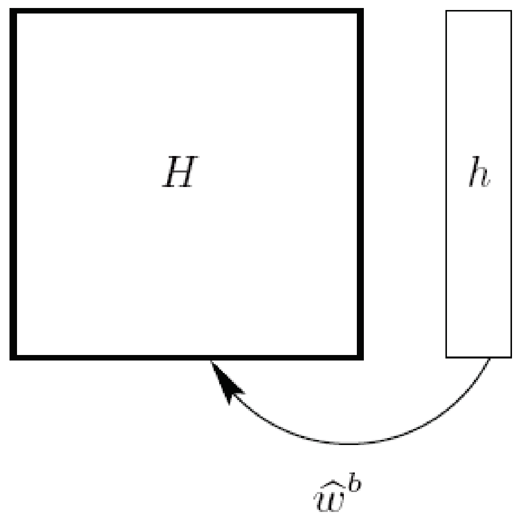

FIGURE 32.1 Backward projection: The last column of $[H \;\; h]$ is projected onto the column span of the matrix H.

BACKWARD UPDATE RELATIONS

We denote the resulting estimate and residual vectors of this projection problem by

$$\begin{aligned}\widehat{h} &= H\widehat{w}^b = HPH^*Wh \\ \widetilde{h} &= h - \widehat{h} = h - H\widehat{w}^b = h - HPH^*Wh\end{aligned}$$

The last entry of $\widetilde{h}$ is $\alpha - u\widehat{w}^b$; it corresponds to the error in estimating α and we denote it by

$$\boxed{\widetilde{\alpha} \triangleq \alpha - u\widehat{w}^b} \tag{32.18}$$

Also, the minimum cost of (32.17) is given by (cf. Table 29.3):

$$\boxed{\xi_h = h^*W\widetilde{h}} \tag{32.19}$$

Using the above definitions of $\{\widehat{w}^b, \widetilde{h}\}$ we can now rewrite (32.15) as

$$\boxed{P_z = \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\sigma + \xi_h}\begin{bmatrix} -\widehat{w}^b \\ 1 \end{bmatrix}\begin{bmatrix} -\widehat{w}^{b*} & 1 \end{bmatrix}} \tag{32.20}$$

BACKWARD UPDATE RELATIONS

If we multiply both sides of (32.20) from the right by

$$\begin{bmatrix} H^* \\ h^* \end{bmatrix} Wy$$

and use the definitions (32.2) and (32.8) for $\{\widehat{w}_z, \widehat{w}\}$ we find that

$$\widehat{w}_z \;=\; \begin{bmatrix} \widehat{w} \\ 0 \end{bmatrix} + \frac{1}{\sigma + \xi_h} \begin{bmatrix} -\widehat{w}^b \\ 1 \end{bmatrix} \widetilde{h}^* W y$$

If we introduce the scalar quantity

$$\boxed{\kappa = \frac{\widetilde{h}^* W y}{\sigma + \xi_h}} \qquad (32.21)$$

then the above relation between $\{\widehat{w}_z, \widehat{w}\}$ becomes

$$\boxed{\widehat{w}_z \;=\; \begin{bmatrix} \widehat{w} \\ 0 \end{bmatrix} + \kappa \begin{bmatrix} -\widehat{w}^b \\ 1 \end{bmatrix} = \begin{bmatrix} \widehat{w} - \kappa \widehat{w}^b \\ \kappa \end{bmatrix}} \qquad (32.22)$$

BACKWARD UPDATE RELATIONS

Likewise, the projections $\{\widehat{y}, \widehat{y}_z\}$ are related via

$$\widehat{y}_z = \begin{bmatrix} H & h \end{bmatrix} \widehat{w}_z = \begin{bmatrix} H & h \end{bmatrix} \begin{bmatrix} \widehat{w} - \kappa\, \widehat{w}^b \\ \kappa \end{bmatrix} = H(\widehat{w} - \kappa \widehat{w}^b) + \kappa h = \widehat{y} + \kappa(h - H\widehat{w}^b)$$

That is,

$$\boxed{\widehat{y}_z = \widehat{y} \; + \; \kappa \widetilde{h}} \qquad (32.23)$$

and, consequently, the residual vectors $\{\widetilde{y}_z, \widetilde{y}\}$ satisfy

$$\boxed{\widetilde{y}_z = \widetilde{y} \; - \; \kappa \widetilde{h}} \qquad (32.24)$$

It is also straightforward to see that the corresponding minimum costs $\{\xi, \xi_z\}$ satisfy

$$\xi_z \;\; = \;\; y^* W \widetilde{y}_z \; = \; y^* W[\widetilde{y} \; - \; \kappa \widetilde{h}] \; = \; y^* W \widetilde{y} \; - \; \kappa y^* W \widetilde{h}$$

i.e., using (32.21),

$$\boxed{\xi_z = \xi - \frac{|\rho|^2}{\sigma + \xi_h}} \qquad (32.25)$$

BACKWARD UPDATE RELATIONS

Finally, if we multiply both sides of (32.20) by $[u \quad \alpha]$ from the left and $\text{col}\{u^*, \alpha^*\}$ from the right, and use the definitions of $\{\gamma_z, \gamma, \widetilde{\alpha}\}$ in (32.6), (32.12), and (32.18), we find that

$$\boxed{\gamma_z = \gamma - \frac{|\widetilde{\alpha}|^2}{\sigma + \xi_h}} \tag{32.27}$$

BACKWARD UPDATE RELATIONS

Lemma 32.1 (Backward order-updates) Consider the regularized least-squares problems (32.1) and (32.7), where the data matrix H in the first problem is extended by one column to $[H \;\; h]$ in the second problem. The corresponding solutions $\{\widehat{w}, \widehat{w}_z\}$, coefficient matrices $\{P, P_z\}$, residuals $\{\widetilde{y}, \widetilde{y}_z\}$, minimum costs $\{\xi, \xi_z\}$, and conversion factors $\{\gamma, \gamma_z\}$ are related as follows.

1. Let $\widehat{w}^b$ be the weight vector that projects h onto H, according to (32.17), i.e., $\widehat{w}^b = PH^*Wh$, where $P = (\Pi + H^*WH)^{-1}$. Let $\widetilde{h}$ denote the resulting residual vector, $\widetilde{h} = h - H\widehat{w}^b$, whose last entry is $\widetilde{\alpha}$. Let also ξ_h denote the corresponding minimum cost, $\xi_h = h^*W\widetilde{h}$.

2. Define the scalars $\rho = y^*W\widetilde{h}$ and $\kappa = \rho^*/(\sigma + \xi_h)$.

Then the following relations hold:

$$
\begin{aligned}
\widehat{y}_z &= \widehat{y} + \kappa\widetilde{h} \\
\widetilde{y}_z &= \widetilde{y} - \kappa\widetilde{h} \\
\xi_z &= \xi - |\rho|^2/(\sigma + \xi_h) \\
\gamma_z &= \gamma - |\widetilde{\alpha}|^2/(\sigma + \xi_h)
\end{aligned}
$$

$$
\begin{aligned}
\widehat{w}_z &= \begin{bmatrix} \widehat{w} \\ 0 \end{bmatrix} + \kappa \begin{bmatrix} -\widehat{w}^b \\ 1 \end{bmatrix} \\
P_z &= \begin{bmatrix} P & 0 \\ 0 & 0 \end{bmatrix} + \frac{1}{\sigma + \xi_h} \begin{bmatrix} -\widehat{w}^b \\ 1 \end{bmatrix} \begin{bmatrix} -\widehat{w}^{b*} & 1 \end{bmatrix}
\end{aligned}
$$

ABSENCE OF REGULARIZATION

In order to illustrate how the presence of regularization affects the interpretation of the results, let us examine what happens when regularization is ignored. Thus assume that $\Pi = 0$ and $\sigma = 0$, so that the backward projection problem (32.17) reduces to

$$\min_{w^b} \quad (h - Hw^b)^* W (h - Hw^b) \tag{32.28}$$

Its solution $\widehat{w}^b$ is now such that it satisfies the orthogonality condition

$$H^* W (h - H\widehat{w}^b) = 0 \tag{32.29}$$

and, in this case, we get $\widehat{h}^* W \widetilde{h} = 0$ since $\widehat{h} \in \mathcal{R}(H)$. This fact allows us to rewrite the minimum cost $\xi_h = h^* W \widetilde{h}$ in (32.19) as

$$\xi_h \;=\; h^* W \widetilde{h} \;=\; [\widehat{h} + \widetilde{h}]^* W \widetilde{h} = \widetilde{h}^* W \widetilde{h} \tag{32.30}$$

and expression (32.21) for κ becomes

$$\kappa = \frac{\widetilde{h}^* W y}{\widetilde{h}^* W \widetilde{h}} \tag{32.31}$$

ABSENCE OF REGULARIZATION

This expression identifies κ as the coefficient that projects y onto $\widetilde{h}$, i.e., it is the coefficient that solves the least-squares problem:

$$\min_k \ (y - \widetilde{h}k)^* W (y - \widetilde{h}k) \quad \Longrightarrow \quad \kappa$$

In other words, when $\Pi = 0$ and $\sigma = 0$, the term $\kappa\widetilde{h}$ in (32.23) can be interpreted as the projection of y onto $\widetilde{h}$. The equality (32.31) does *not* hold in the regularized case and, therefore, we cannot interpret κ in that case as the solution to the problem of projecting y onto $\widetilde{h}$. That is, in the regularized case (32.17), it does *not* hold that κ solves

$$\min_k \ \sigma|k|^2 \ + \ (y - \widetilde{h}k)^* W (y - \widetilde{h}k)$$

since the solution to this problem would be $\kappa_\sigma = \widetilde{h}^* W y / (\sigma + \widetilde{h}^* W \widetilde{h})$, with the term $\widetilde{h}^* W \widetilde{h}$ appearing in the denominator instead of $h^* W \widetilde{h}$, as in expression (32.21) for κ.

ABSENCE OF REGULARIZATION

Actually, more can be said about κ in (32.31). If we write

$$y = \widehat{y} + \widetilde{y}$$

and recall that $\widehat{y} \in \mathcal{R}(H)$, then by virtue of the orthogonality condition (32.29) we have $\widetilde{h}^* W \widehat{y} = 0$, so that

$$\widetilde{h}^* W y \;=\; \widetilde{h}^* W (\widehat{y} + \widetilde{y}) \;=\; \widetilde{h}^* W \widetilde{y}$$

and (32.31) can be replaced by

$$\kappa = \frac{\widetilde{h}^* W \widetilde{y}}{\widetilde{h}^* W \widetilde{h}} \tag{32.32}$$

That is, in the un-regularized case (32.28), we can also interpret κ as the coefficient that projects $\widetilde{y}$ onto $\widetilde{h}$, i.e., it solves

$$\min_k \;\; (\widetilde{y} - \widetilde{h}k)^* W (\widetilde{y} - \widetilde{h}k) \quad \Longrightarrow \quad \kappa \tag{32.33}$$

In view of this result, the difference $\widetilde{y} - \kappa\widetilde{h}$ in (32.24) can be interpreted as the residual vector that results from (32.33).

ABSENCE OF REGULARIZATION

In conclusion, from (32.23)–(32.24) and from (32.31)–(32.32), we have that in the *un-regularized* case the new projection $\widehat{y}_z$ can be obtained from the old projection y as follows:

1. Project h onto $\mathcal{R}(H)$ and find the residual vector $\widetilde{h}$.

2. Project y onto $\widetilde{h}$.

3. Then $\widehat{y}_z = \widehat{y} + \widehat{y}_{\widetilde{h}}$, where $\widehat{y}_{\widetilde{h}}$ denotes the projection of y onto $\widetilde{h}$.

4. Also, $\widetilde{y}_z = \widetilde{y} - \kappa\widetilde{h}$.

That is, we obtain the new projection, $\widehat{y}_z$, by projecting y separately onto H and $\widetilde{h}$ and adding the results. Alternatively, we can obtain the new residual vector $\widetilde{y}_z$ of step 4) as follows:

4.a) Find the residual vector $\widetilde{y}$ that results from projecting y onto $\mathcal{R}(H)$.

4.b) Find the residual vector $\widetilde{h}$ that results from projecting h onto $\mathcal{R}(H)$.

4.c) Then $\widetilde{y}_z$ is the residual vector that results from projecting $\widetilde{y}$ onto $\widetilde{h}$.

ABSENCE OF REGULARIZATION

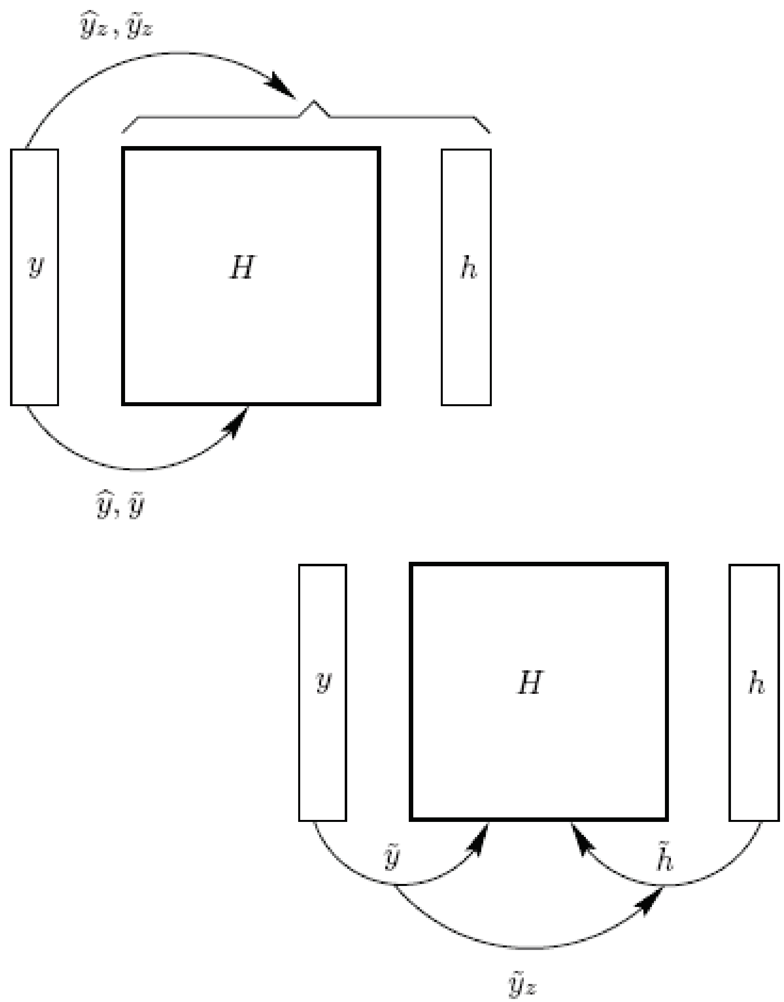

FIGURE 32.2 For un-regularized least-squares problems, the order-update of the residual vector $\widetilde{y}$ into the residual vector $\widetilde{y}_z$ (*top figure*) is obtained by projecting $\widetilde{y}$ onto $\widetilde{h}$ (*bottom figure*).

32.2 FORWARD ORDER-UPDATE RELATIONS

In Sec. 32.1 we started from the least-squares problem (32.1) and extended the data matrix H by adding a column h to its right, as in (32.7). Then in Lemma 32.1 we related the solutions to both problems. We can similarly consider the problem of extending H by adding a column to its *left* (rather than to its right), and by extending Π accordingly, say

$$\min_{w_z} \left\{ w_z^* \begin{bmatrix} \sigma & 0 \\ 0 & \Pi \end{bmatrix} w_z + \left(y - \begin{bmatrix} h & H \end{bmatrix} w_z\right)^* W \left(y - \begin{bmatrix} h & H \end{bmatrix} w_z\right) \right\} \quad (32.39)$$

The optimal solution of this extended problem is now given by

$$\widehat{w}_z = \left(\begin{bmatrix} \sigma & 0 \\ 0 & \Pi \end{bmatrix} + \begin{bmatrix} h^* \\ H^* \end{bmatrix} W \begin{bmatrix} h & H \end{bmatrix} \right)^{-1} \begin{bmatrix} h^* \\ H^* \end{bmatrix} W y$$

We again denote the inverse of its extended coefficient matrix by

$$\boxed{P_z = \left(\begin{bmatrix} \sigma & 0 \\ 0 & \Pi \end{bmatrix} + \begin{bmatrix} h^* \\ H^* \end{bmatrix} W \begin{bmatrix} h & H \end{bmatrix} \right)^{-1}} \quad (32.40)$$

FORWARD UPDATE RELATIONS

so that

$$\widehat{y}_z = \begin{bmatrix} h & H \end{bmatrix} \widehat{w}_z = \begin{bmatrix} h & H \end{bmatrix} P_z \begin{bmatrix} h^* \\ H^* \end{bmatrix} W y$$

The resulting residual vector is

$$\widetilde{y}_z = y - \begin{bmatrix} h & H \end{bmatrix} \widehat{w}_z$$

with the minimum cost of (32.39) equal to (cf. Table 29.3):

$$\xi_z = y^* W \widetilde{y}_z$$

FORWARD UPDATE RELATIONS

We further associate with (32.39) the conversion factor

$$\boxed{\gamma_z = 1 - \begin{bmatrix} \alpha & u \end{bmatrix} P \begin{bmatrix} \alpha^* \\ u^* \end{bmatrix}}$$

where $[\alpha \quad u]$ denotes the last row of $[h \quad H]$.

We can again relate the quantities $\{P, P_z\}$, $\{w, w_z\}$, $\{\widehat{y}, \widehat{y}_z\}$, and $\{\gamma, \gamma_z\}$ of problems (32.1) and (32.39). The arguments (both algebraic and geometric) are identical to those in Sec. 32.1. For this reason, we only state the final results here and leave the details to Probs. VII.15–VII.17.

FORWARD UPDATE RELATIONS

The relation between problems (32.1) and (32.39) is determined in terms of the solution to the following regularized least-squares problem,

$$\min_{w^f} \left[w^{f*}\Pi w^f \; + \; (h - Hw^f)^* W (h - Hw^f) \right] \tag{32.41}$$

namely, $\widehat{w}^f = PH^*Wh$. This weight vector performs the (regularized) projection of h onto the column span of H. And this projection problem is the reason for the title of this subsection, namely, *forward projection*. The terminology refers to the fact that we are projecting h onto the posterior columns of $[h \;\; H]$ — see Fig. 32.3. The minimum cost of (32.41) is given by

$$\boxed{\xi_h \; = \; h^* W \widetilde{h}} \tag{32.42}$$

where

$$\widetilde{h} = h - \widehat{h} = h - H\widehat{w}^f = h - HPH^*Wh$$

FORWARD UPDATE RELATIONS

Lemma 32.2 (Forward order-updates) Consider the regularized problems (32.1) and (32.39), where the data matrix H in the first problem is extended by one column to $[h\ \ H]$ in the second problem. The corresponding solutions $\{\widehat{w}, \widehat{w}_z\}$, coefficient matrices $\{P, P_z\}$, residuals $\{\widetilde{y}, \widetilde{y}_z\}$, minimum costs $\{\xi, \xi_z\}$, and conversion factors $\{\gamma, \gamma_z\}$ are related as follows.

1. Let $\widehat{w}^f$ be the weight vector that projects h onto H, according to (32.41), i.e., $\widehat{w}^f = PH^*Wh$, where $P = (\Pi + H^*WH)^{-1}$. Let $\widetilde{h}$ denote the resulting residual vector, $\widetilde{h} = h - H\widehat{w}^f$, whose last entry is $\widetilde{\alpha}$. Let also ξ_h denote the corresponding minimum cost, $\xi_h = h^*W\widetilde{h}$.
2. Define the scalars $\rho = y^*W\widetilde{h}$ and $\kappa = \rho^*/(\sigma + \xi_h)$.

Then the following relations hold

$$
\begin{aligned}
\widehat{y}_z &= \widehat{y} + \kappa\widetilde{h} \\
\widetilde{y}_z &= \widetilde{y} - \kappa\widetilde{h} \\
\xi_z &= \xi - |\rho|^2/(\sigma + \xi_h) \\
\gamma_z &= \gamma - |\widetilde{\alpha}|^2/(\sigma + \xi_h)
\end{aligned}
$$

$$
\begin{aligned}
\widehat{w}_z &= \begin{bmatrix} 0 \\ \widehat{w} \end{bmatrix} + \kappa \begin{bmatrix} 1 \\ -\widehat{w}^f \end{bmatrix} \\
P_z &= \begin{bmatrix} 0 & 0 \\ 0 & P \end{bmatrix} + \frac{1}{\sigma + \xi_h} \begin{bmatrix} 1 \\ -\widehat{w}^f \end{bmatrix} \begin{bmatrix} 1 & -\widehat{w}^{f*} \end{bmatrix}
\end{aligned}
$$

ABSENCE OF REGULARIZATION

Let us examine again what happens when $\Pi = 0$ and $\sigma = 0$ in problem (32.39), which then reduces to

$$\min_{w^f} \ (h - Hw^f)^* W (h - Hw^f) \tag{32.43}$$

Its solution is such that it should satisfy the orthogonality condition $H^* W (h - H\widehat{w}^f) = 0$. Therefore, in this case, we also have $\widehat{h}^* W \widetilde{h} = 0$ since $\widehat{h} \in \mathcal{R}(H)$. This fact allows us to rewrite the minimum cost $\xi_h = h^* W \widetilde{h}$ in (32.42) as

$$\xi_h = h^* W \widetilde{h} \ = \ [\widehat{h} + \widetilde{h}]^* W \widetilde{h} = \widetilde{h}^* W \widetilde{h} \tag{32.44}$$

so that the expression for κ in Lemma 32.2 becomes

$$\kappa = \frac{\widetilde{h}^* W y}{\widetilde{h}^* W \widetilde{h}} \tag{32.45}$$

As explained in the case of backward projection following Lemma 32.1, the above expression allows us to identify κ as the coefficient that projects y onto $\widetilde{h}$, i.e., it solves the least-squares problem

$$\min_{k} \ (y - \widetilde{h}k)^* W (y - \widetilde{h}k) \quad \Longrightarrow \quad \kappa$$

ABSENCE OF REGULARIZATION

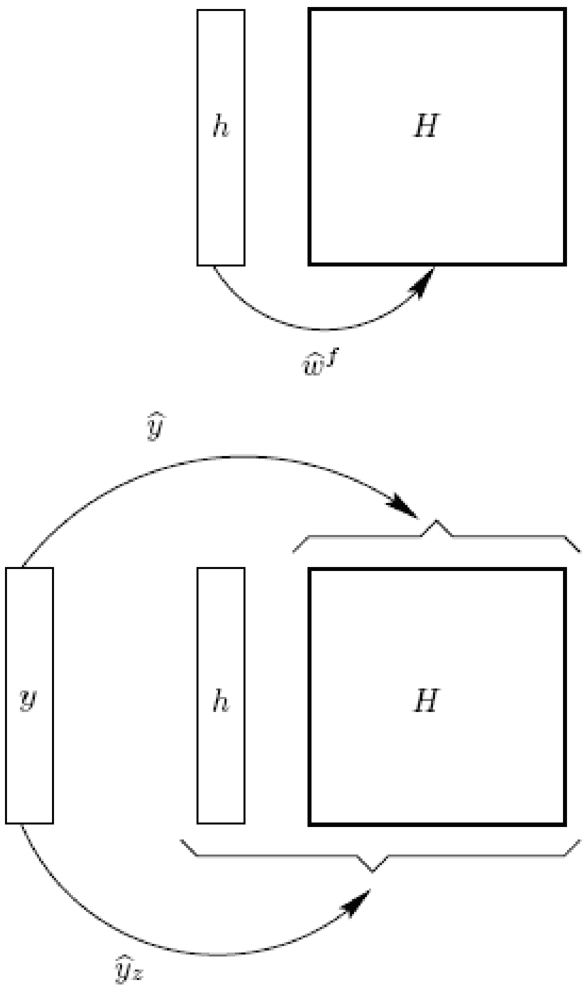

FIGURE 32.3 Forward projection: The leading column of $[h \quad H]$ is projected onto the column span of H (*top figure*). This step helps relate the order-update problems of projecting y onto the column spans of H and $[h \quad H]$ (*bottom figure*).

ABSENCE OF REGULARIZATION

More can be said about κ in (32.45). If we write

$$y = \widehat{y} + \widetilde{y}$$

and recall that $\widehat{y} \in \mathcal{R}(H)$, then since $\widetilde{h}^* W \widehat{y} = 0$ by the orthogonality condition of weighted least-squares solutions, we get

$$\widetilde{h}^* W y = \widetilde{h}^* W(\widehat{y} + \widetilde{y}) = \widetilde{h}^* W \widetilde{y}$$

and (32.45) can be replaced by

$$\kappa = \frac{\widetilde{h}^* W \widetilde{y}}{\widetilde{h}^* W \widetilde{h}} \tag{32.46}$$

That is, in the un-regularized case (32.43), we can also interpret κ as the coefficient that projects $\widetilde{y}$ onto $\widetilde{h}$, i.e., it solves

$$\min_{k} \; (\widetilde{y} - \widetilde{h}k)^* W (\widetilde{y} - \widetilde{h}k) \quad \Longrightarrow \quad \kappa \tag{32.47}$$

In view of this result, the difference $\widetilde{y} - \kappa\widetilde{h}$ in the expression for $\widetilde{y}_z$ in Lemma 32.2 can be interpreted as the residual vector that results from (32.47).

ABSENCE OF REGULARIZATION

In conclusion, we find that in the *un-regularized* case the new projection $\widehat{y}_z$ can be obtained from the old projection y as follows:

1. Project h onto $\mathcal{R}(H)$ and find the residual vector $\widetilde{h}$.

2. Project y onto $\widetilde{h}$.

3. Then $\widehat{y}_z = \widehat{y} + \widehat{y}_{\widetilde{h}}$, where $\widehat{y}_{\widetilde{h}}$ denotes the projection of y onto $\widetilde{h}$.

4. Also, $\widetilde{y}_z = \widetilde{y} - \kappa\widetilde{h}$.

That is, we obtain the new projection, $\widehat{y}_z$, by projecting y separately onto H and $\widetilde{h}$ and adding the results. Alternatively, we can obtain the new residual vector $\widetilde{y}_z$ of step 4) as follows:

4.a) Find the residual vector $\widetilde{y}$ that results from projecting y onto $\mathcal{R}(H)$.

4.b) Find the residual vector $\widetilde{h}$ that results from projecting h onto $\mathcal{R}(H)$.

4.c) Then $\widetilde{y}_z$ is the residual vector that results from projecting $\widetilde{y}$ onto $\widetilde{h}$.

This construction of $\widetilde{y}_z$ in the un-regularized case is depicted in Fig. 32.4. In the figure, the arrows indicate that $\widetilde{y}$ results from projecting y onto $\mathcal{R}(H)$ and $\widetilde{h}$ results from projecting h onto $\mathcal{R}(H)$.

ABSENCE OF REGULARIZATION

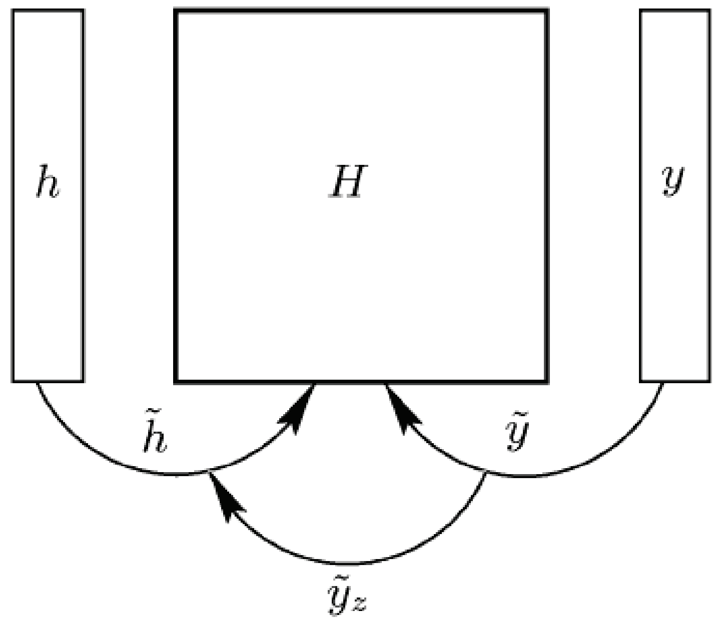

FIGURE 32.4 For un-regularized least-squares problems, the order-update of the residual vector $\tilde{y}$ into the residual vector $\tilde{y}_z$ is obtained by projecting $\tilde{y}$ onto $\widetilde{h}$.

32.3 TIME-UPDATE RELATION

The discussion in this section complements the one presented in Secs. 32.1 and 32.2. The results derived here will only be used later in Part X (*Lattice Filters*), when we study order-recursive adaptive filters.

Thus consider an $(N-1) \times M$ data matrix H_{N-1} and partition it as

$$H_{N-1} \triangleq \left[\begin{array}{ccc} x_{N-1} & \bar{H}_{N-1} & z_{N-1} \end{array} \right] \tag{32.48}$$

where x_{N-1} and z_{N-1} denote the leading and trailing columns of H_{N-1}, and $\bar{H}_{N-1}$ denotes its middle columns. Let $\widehat{z}_{N-1}$ denote the (regularized) projection of z_{N-1} onto $\mathcal{R}(\bar{H}_{N-1})$, i.e.,

$$\widehat{z}_{N-1} = \bar{H}_{N-1} w_{N-1,z}$$

TIME-UPDATE RELATION

where $w_{N-1,z}$ is obtained by solving the regularized least-squares problem

$$\min_{w} \left[\lambda^N w^* \Pi w + (z_{N-1} - \bar{H}_{N-1} w)^* \Lambda_{N-1} (z_{N-1} - \bar{H}_{N-1} w) \right] \tag{32.49}$$

and Λ_{N-1} is defined as in (30.25). Let further $\widetilde{z}_{N-1}$ denote the resulting residual vector,

$$\widetilde{z}_{N-1} = z_{N-1} - \widehat{z}_{N-1} \; = \; z_{N-1} - \bar{H}_{N-1} w_{N-1,z}$$

and define the weighted inner product

$$\boxed{\Delta(N-1) \; \stackrel{\Delta}{=} \; x_{N-1}^* \Lambda_{N-1} \widetilde{z}_{N-1}}$$

In other words, $\Delta(N-1)$ is the weighted inner product between the first column of H_{N-1} and the residual $\widetilde{z}_{N-1}$ that results from projecting its last column onto the middle columns $\bar{H}_{N-1}$ — see Fig. 32.5.

TIME-UPDATE RELATION

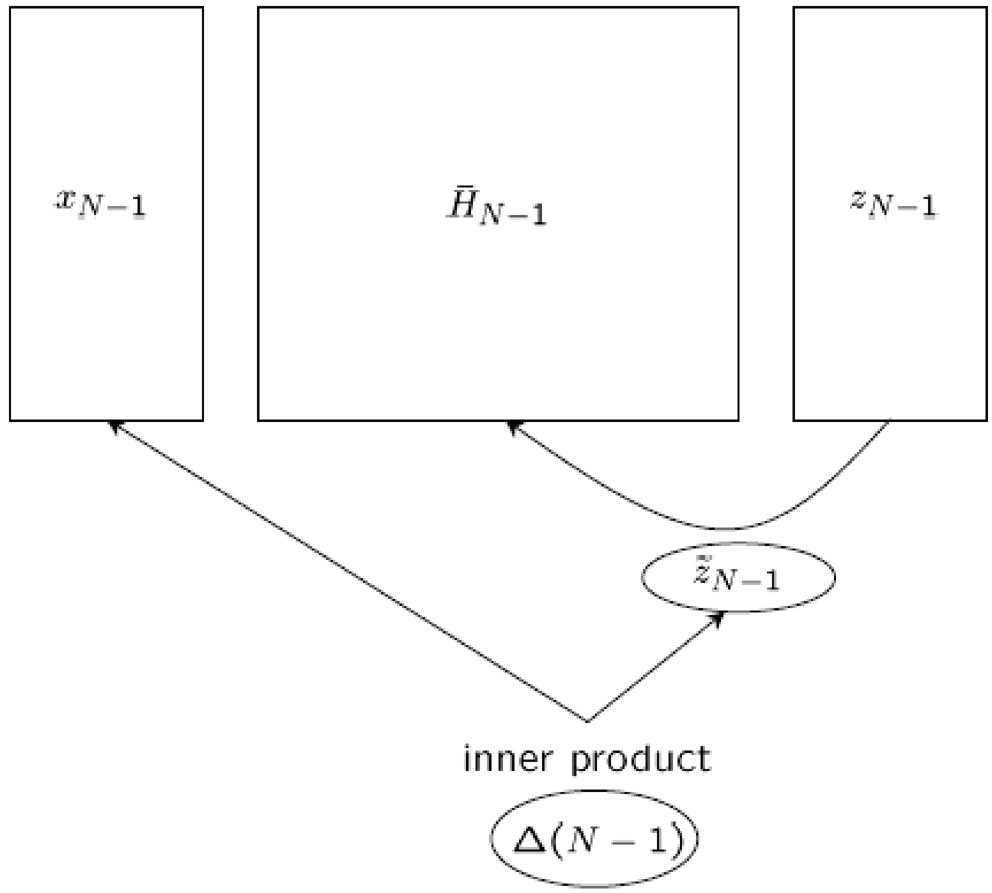

FIGURE 32.5 Inner product between the residual vector $\tilde{z}_{N-1}$ and x_{N-1}.

TIME-UPDATE RELATION

Now assume that one more row is appended to H_{N-1} in (32.48), say

$$H_N \triangleq \begin{bmatrix} x_{N-1} & \bar{H}_{N-1} & z_{N-1} \\ \alpha(N) & \bar{h}_N & \beta(N) \end{bmatrix} \triangleq \begin{bmatrix} x_N & \bar{H}_N & z_N \end{bmatrix} \tag{32.50}$$

where $\alpha(N)$ and $\beta(N)$ are scalars, while $\bar{h}_N$ is a row vector. That is, H_{N-1} is *time-updated* to H_N.

As above, let $\widehat{z}_N$ denote the (regularized) projection of z_N onto $\mathcal{R}(\bar{H}_N)$, i.e.,

$$\widehat{z}_N = \bar{H}_N w_{N,z}$$

where $w_{N,z}$ is obtained by solving the regularized least-squares problem

$$\min_w \left[\lambda^{(N+1)} w^* \Pi w \; + \; (z_N - \bar{H}_N w)^* \Lambda_N (z_N - \bar{H}_N w) \right] \tag{32.51}$$

TIME-UPDATE RELATION

Let further $\widetilde{z}_N$ denote the resulting residual vector,

$$\widetilde{z}_N = z_N - \widehat{z}_N = z_N - \bar{H}_N w_{N,z}$$

and define the corresponding weighted inner product

$$\boxed{\Delta(N) \stackrel{\Delta}{=} x_N^* \Lambda_N \widetilde{z}_N} \qquad (32.52)$$

Again, $\Delta(N)$ is the weighted inner product between the first column of H_N and the residual vector that results from projecting its last column onto the middle columns $\bar{H}_N$ — see Fig. 32.6.

TIME-UPDATE RELATION

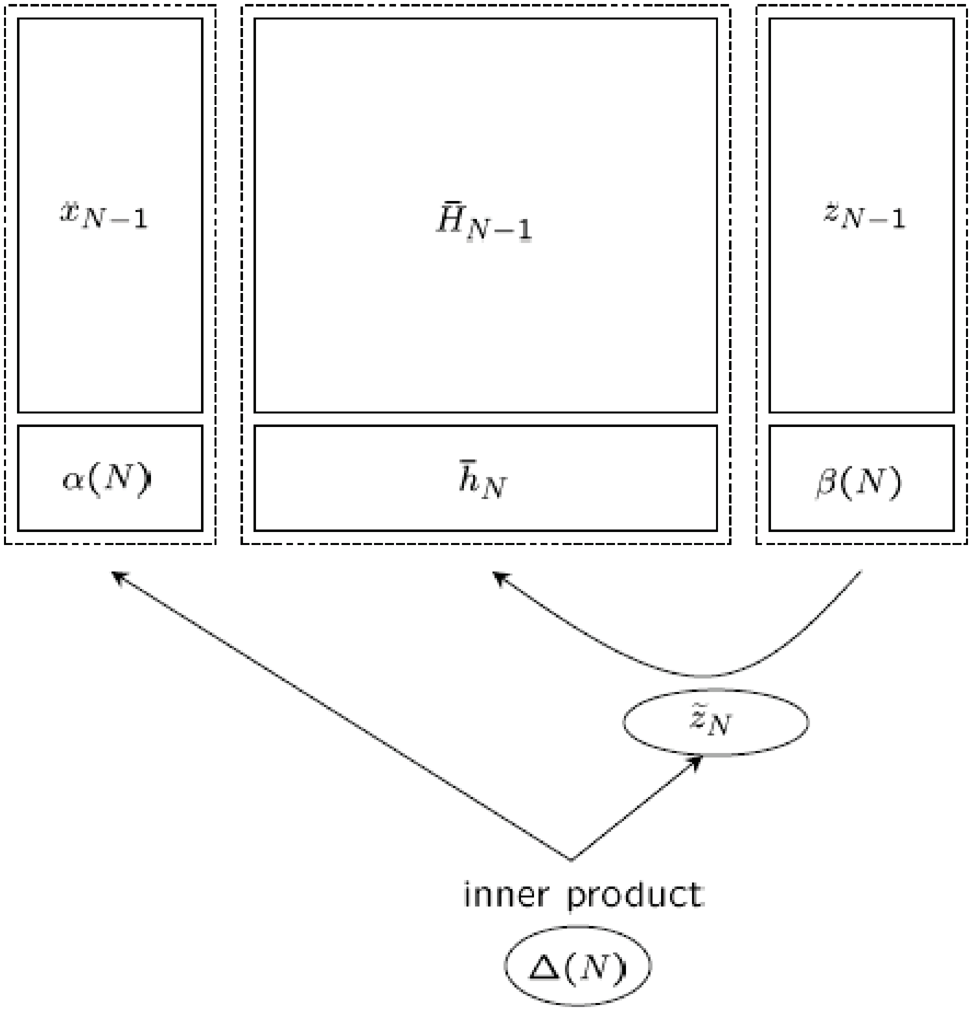

FIGURE 32.6 Inner product between the residual vector $\tilde{z}_N$ and x_N.

TIME-UPDATE RELATION

We would like to relate $\Delta(N)$ and $\Delta(N-1)$, i.e., we would like to determine a time-update relation for the variable Δ. For this purpose, let $\widehat{x}_N$ denote the (regularized) projection of x_N onto $\mathcal{R}(\bar{H}_N)$:

$$\widehat{x}_N = \bar{H}_N w_{N,x}$$

where $w_{N,x}$ is obtained by solving the regularized least-squares problem

$$\min_w \left[\lambda^{(N+1)} w^* \Pi w \ + \ (x_N - \bar{H}_N w)^* \Lambda_N (x_N - \bar{H}_N w) \right] \tag{32.53}$$

Introduce the estimation errors

$$\boxed{\begin{array}{rcl} \widetilde{\alpha}(N) & = & \alpha(N) - \bar{h}_N w_{N,x} \\ \widetilde{\beta}(N) & = & \beta(N) - \bar{h}_N w_{N,z} \end{array}} \tag{32.54}$$

Here, $\widetilde{\alpha}(N)$ is the *a posteriori* error in estimating the last entry of x_N, while $\widetilde{\beta}(N)$ is the *a posteriori* error in estimating the last entry of z_N.

TIME-UPDATE RELATION

Define further the conversion factor

$$\boxed{\bar{\gamma}(N) \;\stackrel{\Delta}{=}\; 1 - \bar{h}_N \bar{P}_N \bar{h}_N^*} \qquad (32.55)$$

where

$$\bar{P}_N \;=\; \left[\lambda^{(N+1)}\Pi + \bar{H}_N^* \Lambda_N \bar{H}_N\right]^{-1}$$

Clearly, $\bar{\gamma}(N)$ is the factor that relates the *a posteriori* error $\widetilde{\beta}(N)$ to its *a priori* version, which is defined by

$$\widetilde{\beta}_a(N) \;=\; \beta(N) - \bar{h}_N w_{N-1,z}$$

with $w_{N-1,z}$ used instead of $w_{N,z}$. That is, $\widetilde{\beta}(N) = \bar{\gamma}(N)\widetilde{\beta}_a(N)$. Actually, $\bar{\gamma}(N)$ is also the factor that relates the *a posteriori* error $\widetilde{\alpha}(N)$ to its *a priori* version, which is defined as

$$\widetilde{\alpha}_a(N) \;=\; \alpha(N) - \bar{h}_N w_{N-1,x}$$

with $w_{N-1,x}$ used instead of $w_{N,x}$, and where $w_{N-1,x}$ is the solution to a problem similar to (32.49) with z_{N-1} replaced by x_{N-1}. That is, $\widetilde{\alpha}(N) = \bar{\gamma}(N)\widetilde{\alpha}_a(N)$.

TIME-UPDATE RELATION

Now from the definition (32.52) for $\Delta(N)$, and using the fact that

$$\Lambda_N = \begin{bmatrix} \lambda\Lambda_{N-1} & 0 \\ 0 & 1 \end{bmatrix}$$

we obtain

$$\begin{aligned}
\Delta(N) &= \begin{bmatrix} x^*_{N-1} & \alpha^*(N) \end{bmatrix} \begin{bmatrix} \lambda\Lambda_{N-1} & 0 \\ 0 & 1 \end{bmatrix} \left(\begin{bmatrix} z_{N-1} \\ \beta(N) \end{bmatrix} - \begin{bmatrix} \bar{H}_{N-1} \\ \bar{h}_N \end{bmatrix} w_{N,z} \right) \\
&= \lambda x^*_{N-1}\Lambda_{N-1}z_{N-1} + \alpha^*(N)\beta(N) - (\lambda x^*_{N-1}\Lambda_{N-1}\bar{H}_{N-1} + \alpha^*(N)\bar{h}_N)w_{N,z} \\
&= \lambda x^*_{N-1}\Lambda_{N-1}z_{N-1} + \alpha^*(N)\widetilde{\beta}(N) - \lambda x^*_{N-1}\Lambda_{N-1}\bar{H}_{N-1}w_{N,z}
\end{aligned}$$

Moreover, the RLS recursion (30.13) allows us to relate $w_{N,z}$ and $w_{N-1,z}$ as

$$w_{N,z} = w_{N-1,z} + \bar{P}_N\bar{h}^*_N\left[\widetilde{\beta}(N)/\bar{\gamma}(N)\right]$$

TIME-UPDATE RELATION

Substituting into the expression for $\Delta(N)$, and using

$$\lambda\bar{\gamma}^{-1}(N)\bar{P}_N\bar{h}_N^* = \bar{P}_{N-1}\bar{h}_N^*$$

we obtain, after grouping terms,

$$\boxed{\Delta(N) = \lambda\Delta(N-1) + \widetilde{\alpha}^*(N)\widetilde{\beta}(N)/\bar{\gamma}(N)} \quad (32.56)$$

This is a useful relation and it plays an important role in the derivation of order-recursive algorithms (see Part X (*Lattice Filters*)). The relation holds for generic regularized least-squares problems and there are no structural restrictions imposed on the data matrices $\{H_{N-1}, H_N\}$.

TIME-UPDATE RELATIONS

Lemma 32.3 (Inner product time-update) Consider the data matrix (32.50), namely,

$$H_N \triangleq \begin{bmatrix} x_{N-1} & \bar{H}_{N-1} & z_{N-1} \\ \alpha(N) & \bar{h}_N & \beta(N) \end{bmatrix} = \begin{bmatrix} x_N & \bar{H}_N & z_N \end{bmatrix}$$

Let $\widetilde{z}_{N-1}$ and $\widetilde{z}_N$ denote the residual vectors that result from projecting z_{N-1} onto $\mathcal{R}(\bar{H}_{N-1})$ and z_N onto $\mathcal{R}(\bar{H}_N)$; both projections are meant in the regularized least-squares senses (32.49) and (32.53). Then the weighted inner products

$$\Delta(N) = x_N^* \Lambda_N \widetilde{z}_N, \qquad \Delta(N-1) = x_{N-1}^* \Lambda_{N-1} \widetilde{z}_{N-1}$$

are related via the time-update relation

$$\Delta(N) = \lambda \Delta(N-1) + \widetilde{\alpha}^*(N) \widetilde{\beta}(N) / \bar{\gamma}(N)$$

where $\{\widetilde{\alpha}(N), \widetilde{\beta}(N)\}$ are the *a posteriori* errors in estimating the entries $\{\alpha(N), \beta(N)\}$, as defined by (32.54), and $\bar{\gamma}(N)$ is the conversion factor defined by (32.55).