where the projection matrix $P(U)$ projects $\mathbf{y}$ orthogonally on to the plane formed by all linear combinations of the columns of U. Hence the error $\mathbf{y} - P(U)\mathbf{y}$ is orthogonal to each column of U, as is easily verified:

$$U^T(\mathbf{y} - P(U)\mathbf{y}) = U^T\mathbf{y} - U^T\mathbf{y} = \mathbf{0} \tag{8.4.13}$$

By the same token $\mathbf{y} - P([U_p V_q])\mathbf{y}$ is orthogonal to each column of U_p and V_q so, with M denoting the inverse of the normal matrix as before,

$$\begin{bmatrix} U_p^T \\ V_q^T \end{bmatrix}\mathbf{y} - \begin{bmatrix} U_p^T \\ V_q^T \end{bmatrix}[U_p \quad V_q]M_{p+q}[U_p \quad V_q]\mathbf{y} = \mathbf{0} \tag{8.4.14}$$

or in terms of the partitions $M^{(11)}$, $M^{(12)}$, $M^{(21)}$ and $M^{(22)}$ of M_{p+q},

$$U_p^T\mathbf{y} - U_p^T(U_pM^{(11)} + V_qM^{(21)})U_p^T\mathbf{y} - U_p^T(U_pM^{(12)} + V_qM^{(22)})V_q^T\mathbf{y} = \mathbf{0} \tag{8.4.15}$$

$$V_q^T\mathbf{y} - V_q^T(U_pM^{(11)} + V_qM^{(21)})U_p^T\mathbf{y} - V_q^T(U_pM^{(12)} + V_qM^{(22)})V_q^T\mathbf{y} = \mathbf{0} \tag{8.4.16}$$

Now (8.4.15) must be true for any values of U_p, V_q and $\mathbf{y}$, including $\mathbf{y}$ non-zero but $U_p^T\mathbf{y}$ zero, so that

$$U_p^T(U_pM^{(12)} + V_qM^{(22)})V_q^T\mathbf{y} = \mathbf{0} \tag{8.4.17}$$

To satisfy (8.4.17) it is sufficient to make $U_p^T(U_pM^{(12)} + V_qM^{(22)})$ zero, giving

$$M^{(12)} = -M_pU_p^TV_qM^{(22)} \tag{8.4.18}$$

Similarly, (8.4.16) requires that

$$V_q^T\mathbf{y} - V_q^T(U_pM^{(12)} + V_qM^{(22)})V_q^T\mathbf{y} = \mathbf{0} \tag{8.4.19}$$

and it is enough if $I - V_q^T(U_pM^{(12)} + V_qM^{(22)})$ is zero, which on substitution of $M^{(12)}$ from (8.4.18) gives

$$M^{(22)} = (V_q^TV_q - V_q^TU_pM_pU_p^TV_q)^{-1} \equiv W \tag{8.4.20}$$

Swapping U_p and V_q all through gives $M^{(11)}$ as in (8.4.8), and $M^{(21)}$ is $M^{(12)T}$, as by definition M_{p+q} is symmetric.

The order-incrementing equations for $\hat{\theta}$ and $\hat{\phi}$ also comes from (8.4.14), but written as

$$\begin{bmatrix} U_p^T \\ V_q^T \end{bmatrix}[\mathbf{y} - U_p\hat{\theta}_{p+q} - V_q\hat{\phi}_{p+q}] = \mathbf{0} \tag{8.4.21}$$

We recognise this immediately as the normal equations (8 4.2).

8.4.3 Lattice Algorithms and Identification

A great deal of interest has been aroused in signal processing by the development of lattice algorithms (Friedlander, 1982). The algorithms employ an a.r. time-series model for l.s. signal estimation, and are implemented as a cascade of identical sections, each corresponding to an increase of one in the a.r. order. The algorithms are attractive for their computational economy and good numerical properties, and are potentially useful for identification. However, their economy depends on the model being an autoregression. For an a.r., the regressor vector at time t is that at $t-1$ shifted down one place and with one new entry at the top. The normal matrix is correspondingly updated mainly by shifting south-east. Without going into the details, we can appreciate that this simplifies a combined time-updating and order-incrementing algorithm greatly. For identification, we are rarely happy with a purely a.r. model, and almost always require an a.r.m.a. model with exogenous inputs plus a noise model, and perhaps also a constant term. The updating is much less simple, with several new samples entering at each update. The result is that computational economy is lost (Robins and Wellstead, 1981), and the lattice method has no overwhelming advantage to counterbalance its relatively complicated programming and difficulty of interpretation in identification.

8.5 MODEL REDUCTION

The fewer parameters a model has, the easier it is to understand and apply. The neatest way to ensure that a model has no more parameters than necessary is to conduct order tests during identification, as described in Chapter 9. Nevertheless, we sometimes have to reduce an existing model, perhaps to check whether order reduction alters the overall behaviour significantly. There are many approaches (Bosley and Lees, 1972) of which we shall examine a few of the most popular, applied to transfer-function rather than state-space models.

8.5.1 Moment Matching: Padé Approximation

One way to fit a reduced transfer function

$$\frac{B_m(s)}{1 + A_m(s)} = \frac{b_0 + b_1s + \cdots + b_{m-1}s^{m-1}}{1 + a_1s + \cdots + a_ms^m} \tag{8.5.1}$$

to a larger continuous-time model is to expand the transfer function of the larger model as a power series in s:

$$H(s) = h_0 + h_1 s + h_2 s^2 + \cdots \infty \tag{8.5.2}$$

then pick the $2m$ coefficients in (8.5.1) so as to match terms of (8.5.2) up to $h_{2m-1}s^{2m-1}$. The process of rational-function approximation via a Taylor series is called *Padé approximation* (Watson, 1980). The numerator and denominator degrees can be chosen at will; we have made $B_m(s)$ of degree one less than $1 + A_m(s)$ to give a realistic finite bandwidth.

Matching h_0 matches the steady-state gain, i.e. the final value $\lim_{s\to 0} sH(s)/s$ of the step response, and we can interpret matching higher powers of s as paying attention to the response to higher derivatives of the input. The significance of the matching is best seen in terms of the impulse response. For a stable system

$$i!\,h_i = \lim_{s\to 0} s\cdot\frac{1}{s}\frac{d^iH(s)}{ds^i} = \lim_{t\to\infty} \mathcal{L}^{-1}\left\{\frac{1}{s}\frac{d^iH(s)}{ds^i}\right\} \tag{8.5.3}$$

and

$$\mathcal{L}^{-1}\left\{\frac{1}{s}\frac{d^iH(s)}{ds^i}\right\} = \int_0^t (-t)^i h(t)\,dt \tag{8.5.4}$$

(Gabel and Roberts, 1980). We see that matching h_i matches the ith time moment of the impulse response $h(t)$. That seems sensible enough.

Example 8.5.1 The model

$$Y(s) = \frac{(1+0.5s)(1+0.2s)}{(1+s)(1+0.25s)(1+0.1s)(1+0.05s)}U(s) + \text{noise}$$

is to be reduced to first or second order. To do so, we write the transfer function numerator and denominator in ascending powers of s then expand $H(s)$ by long division, quicker than repeated differentiation and (8.5.3). We obtain

$$H(s) = 1 - 0.7s + 0.6375s^2 - 0.6265s^3 + \cdots$$

An mth-order reduced model then matches the coefficients of s^0 to s^{2m-1} in

$$B_m(s) = H(s)(1 + A_m(s))$$

(i) First-order model: $b_0 = h_0$ and $b_1 = h_0 a_1 + h_1 = 0$ so $b_0 = 1$, $a_1 = 0.7$

(ii) Second-order model: $b_0 = h_0$, $b_1 = h_0 a_1 + h_1$, $b_2 = h_0 a_2 + h_1 a_1 + h_2 = 0$, $b_3 = h_1 a_2 + h_2 a_1 + h_3 = 0$ so $b_0 = 1$, $b_1 = 0.522$, $a_1 = 1.222$, $a_2 = 0.218$.

Figure 8.5.1 gives the step and impulse responses of the original and

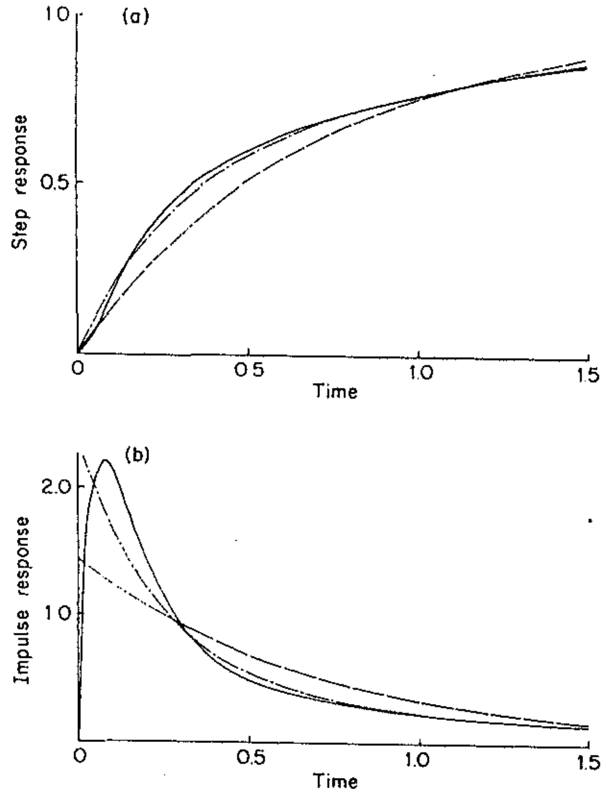

Fig. 8.5.1 (a) Step and (b) impulse responses, Example 8.5.1. ———: original model; — — —: second-order reduced model; —·—: first-order reduced model.

reduced models. Except very early on, the second-order model fits the step response well. The impulse-response fit is less impressive, with quite wrong behaviour initially. The trouble can be traced to a difference in pole-zero excess, two for the original model and one for each reduced model, invalidating the approximation

$$h(0+) = \lim_{s\to\infty} sH(s) \simeq \lim_{s\to\infty} s\frac{B_m(s)}{1 + A_m(s)} \qquad \triangle$$

8.5.2 Continued-Fraction Approximation

A model-reduction method popularised by Chen and Shieh (1968) is to expand the original rational transfer function as a continued fraction in the second