Estimation of a Two Dimensional Correlation Matrix (An Example)

by John Parker Burg

In the following, the covariance estimation principles discussed in the first Stanford Exploration Report will be applied to a simple two dimensional problem. It is assumed that a 2 by 2 array receives N samples from a 2-D stationary process, i.e., the nth sample has the spatial arrangement of

$$\begin{matrix} x_{1n} & x_{2n} \\ x_{3n} & x_{4n} \end{matrix}$$

and the 2-D correlation of the process is

$$\begin{matrix} \phi_{-1\,1} & \phi_{0\,1} & \phi_{1\,1} \\ \phi_{-1\,0} & \phi_{0\,0} & \phi_{1\,0} \\ \phi_{-1-1} & \phi_{0\,-1} & \phi_{1\,-1} \end{matrix} \quad .$$

We first form the 4 by 4 raw covariance matrix from the data by

$$\frac{1}{N}\sum_{n=1}^{N} \begin{Bmatrix} x_{1n} \\ x_{2n} \\ x_{3n} \\ x_{4n} \end{Bmatrix} \left\{ x_{1n} \quad x_{2n} \quad x_{3n} \quad x_{4n} \right\} = \begin{bmatrix} A & E & G & J \\ E & B & I & H \\ G & I & C & F \\ J & H & F & D \end{bmatrix}$$

which is an approximation of the 2-D autocorrelation matrix

$$\begin{bmatrix} \phi_{0\,0} & \phi_{1\,0} & \phi_{0\,1} & \phi_{-1\,1} \\ \phi_{1\,0} & \phi_{0\,0} & \phi_{1\,1} & \phi_{0\,1} \\ \phi_{0\,1} & \phi_{1\,1} & \phi_{0\,0} & \phi_{1\,0} \\ \phi_{-1\,1} & \phi_{0\,1} & \phi_{1\,0} & \phi_{0\,0} \end{bmatrix} \qquad (2)$$

Of course there is no necessity for (1) to have the correct form, i.e., $A = B = C = D$, $E = F$ and $G = H$. Thus, we will change (1) so that it does have this form.

To begin, we estimate $\phi_{0\,0}$ by $(A+B+C+D) / 4$. We next estimate $\phi_{1\,0}$ indirectly by estimating the best coefficient, s , to predict x_1 from x_2 , x_2 from x_1 , x_3 from x_4 and x_4 from x_3 , i.e,, we minimize

$$\sum_{n=1}^{N} [(x_{1n} - s\,x_{2n})^2 + (x_{2n} - s\,x_{1n})^2 + (x_{3n} - s\,x_{4n})^2 + (x_{4n} - s\,x_{3n})^2]$$

which gives us

$$s = \frac{2E + 2F}{A+B+C+D} \quad . \qquad (3)$$

We can likewise estimate the coefficient, t , for (x_1, x_3) and (x_2, x_4) and get

$$t = \frac{2G + 2H}{A+B+C+D} \quad . \qquad (4)$$

At this point, the correlation matrix which we are building has the form

$$\left(\frac{A+B+C+D}{4}\right)\begin{bmatrix} 1 & \frac{2E+2F}{A+B+C+D} & \frac{2G+2H}{A+B+C+D} & r \\ \frac{2E+2F}{A+B+C+D} & 1 & q & \frac{2G+2H}{A+B+C+D} \\ \frac{2G+2H}{A+B+C+D} & q & 1 & \frac{2E+2F}{A+B+C+D} \\ r & \frac{2G+2H}{A+B+C+D} & \frac{2E+2F}{A+B+C+D} & 1 \end{bmatrix} \qquad (5)$$

Also, our estimates so far have been between pairs of variables whose correlations to any third variable have not been specified.

We will now determine the value of q . In this case the correlations of x_2 and x_3 with both x_1 and x_4 have been specified and must be removed. The residuals, normalized to unity, in predicting x_2 and x_3 from x_1 are

$$\varepsilon_2 = \frac{x_2 - s\,x_1}{\sqrt{1-s^2}} \quad \text{and} \quad \varepsilon_3 = \frac{x_3 - t\,x_1}{\sqrt{1-t^2}} \quad .$$

Likewise the normalized residuals in predicting x_2 and x_3 from x_4 are

$$\delta_2 = \frac{x_2 - t\,x_4}{\sqrt{1-t^2}} \quad \text{and} \quad \delta_3 = \frac{x_3 - s\,x_4}{\sqrt{1-s^2}} \quad .$$

From symmetry, we see that the same coefficient q should be used to cross predict $(\varepsilon_2, \varepsilon_3)$ and (δ_2, δ_3) . In fact, from symmetry, we note that the expected values of the two matrices

$$\begin{Bmatrix} x_1 \\ x_2 \\ x_3 \end{Bmatrix} \begin{Bmatrix} x_1 & x_2 & x_3 \end{Bmatrix} \quad \text{and} \quad \begin{Bmatrix} x_4 \\ x_3 \\ x_2 \end{Bmatrix} \begin{Bmatrix} x_4 & x_3 & x_2 \end{Bmatrix} \qquad (6)$$

are the same and should be transformed into

$$\overline{x^2} \ast \begin{bmatrix} 1 & s & t \\ s & 1 & q \\ t & q & 1 \end{bmatrix} . \tag{7}$$

Summing our sample covariance matrices corresponding to (6), we have

$$\begin{bmatrix} A+D & E+F & G+H \\ E+F & B+C & 2\ I \\ G+H & 2\ I & C+B \end{bmatrix} . \tag{8}$$

To get q, we shall operate simultaneously on (7) and (8) to get the 2 by 2 correlation matrix of the normalized residuals and then compare results. Post multiplying by

$$\begin{bmatrix} \frac{-\ s}{\sqrt{1-s^2}} & \frac{-\ t}{\sqrt{1-t^2}} \\ \frac{1}{\sqrt{1-s^2}} & 0 \\ 0 & \frac{1}{\sqrt{1-t^2}} \end{bmatrix}$$

we get from (7)

$$\begin{bmatrix} 0 & 0 \\ \frac{1-s^2}{\sqrt{1-s^2}} & \frac{q-s\,t}{\sqrt{1-t^2}} \\ \frac{q-s\,t}{\sqrt{1-s^2}} & \frac{1-t^2}{\sqrt{1-t^2}} \end{bmatrix} \tag{7'}$$

and from (8)

$$\begin{bmatrix} \dfrac{-s\,(A+D) + (E+F)}{\sqrt{1-s^2}} & \dfrac{-t\,(A+D) + (G+H)}{\sqrt{1-t^2}} \\ \dfrac{-s\,(E+F) + (B+C)}{\sqrt{1-s^2}} & \dfrac{-t\,(E+F) + 2I}{\sqrt{1-t^2}} \\ \dfrac{-s\,(G+H) + 2I}{\sqrt{1-s^2}} & \dfrac{-t\,(G+H) + (C+B)}{\sqrt{1-t^2}} \end{bmatrix} \cdot \qquad (8')$$

Premultiplying by

$$\begin{bmatrix} \dfrac{-s}{\sqrt{1-s^2}} & \dfrac{1}{\sqrt{1-s^2}} & 0 \\ \dfrac{-t}{\sqrt{1-t^2}} & 0 & \dfrac{1}{\sqrt{1-t^2}} \end{bmatrix}$$

we get from (7')

$$\begin{bmatrix} 1 & \dfrac{q-st}{\sqrt{(1-s^2)(1-t^2)}} \\ \dfrac{q-st}{\sqrt{(1-s^2)(1-t^2)}} & 1 \end{bmatrix} \qquad (7'')$$

and from (8')

$$\begin{bmatrix} \dfrac{s^2(A+D) - 2s(E+F) + (B+C)}{1 - s^2} & \dfrac{s\,t\,(A+D) - s(G+H) - t(E+F) + 2I}{\sqrt{(1-s^2)(1-t^2)}} \\ \dfrac{s\,t\,(A+D) - S(G+H) - t(E+F) + 2I}{\sqrt{(1-s^2)(1-t^2)}} & \dfrac{t^2(A+D) - 2t(G+H) + (B+C)}{1 - t^2} \end{bmatrix} . \qquad (8'')$$

We can first note that the upper left term in (8") can be changed by using (3) to replace $2E + 2F$.

$$\frac{s^2(A+D) - s(2E + 2F) + (B+C)}{1 - s^2} = \frac{s^2(A+D) - s^2(A+B+C+D) + (B+C)}{1 - s^2} = B + C .$$

Likewise, the lower right term becomes $B+C$. Thus, using our general principle, the correlation coefficient from (8") is

$$\frac{s\,t\,(A+D) - s(G+H) - t(E+F) + 2\,I}{(B+C)\ \sqrt{(1-s^2)(1-t^2)}}$$

and from (7") it is simply

$$\frac{q - s\,t}{\sqrt{(1-s^2)(1-t^2)}} .$$

Comparing, we see that we want

$$q = s\,t + \frac{s\,t\,(A+D) - s(G+H) - t(E+F) + 2\,I}{B + C} .$$

Using (3) to replace (E+F) and (4) to replace (G+H) , we get

$$q = \frac{st(A+B+C+D) - \frac{1}{2}\, st(A+B+C+D) - \frac{1}{2}\, st(A+B+C+D) + 2\,I}{B + C}$$

or simply

$$q = \frac{2\,I}{B+C} \quad . \tag{9}$$

Surprisingly, this is the same answer as would be obtained if x_1 and x_4 were completely ignored.

We now turn to the problem of getting r . Here, we need to remove both x_2 and x_3 from x_1 and x_4 . To predict x_1 from x_2 and x_3 , we will use an unnormalized prediction error filter given by

$$\begin{bmatrix} 1 & s & t \\ s & 1 & q \\ t & q & 1 \end{bmatrix} \begin{Bmatrix} 1 - q^2 \\ tq - s \\ sq - t \end{Bmatrix} = \begin{Bmatrix} 1 + 2\,stq - q^2 - s^2 - t^2 \\ 0 \\ 0 \end{Bmatrix} \tag{10}$$

Dividing the filter coefficients by $1 - q^2$ would normalize the weight on x_1 to unity, but this will not be necessary to achieve our purposes. Our full matrix that we are trying to get is

$$\begin{bmatrix} 1 & s & t & r \\ s & 1 & q & t \\ t & q & 1 & s \\ r & t & s & 1 \end{bmatrix} \quad . \tag{11}$$

Because of symmetry, we can reverse x_1, x_2, x_3, x_4 and sum (1) to get

$$\begin{bmatrix} A+D & E+F & G+H & 2J \\ E+F & B+C & 2I & G+H \\ G+H & 2I & B+C & E+F \\ 2J & G+H & E+F & A+D \end{bmatrix} = \begin{bmatrix} A' & E' & G' & J' \\ E' & B' & I' & G' \\ G' & I' & B' & E' \\ J' & G' & E' & A' \end{bmatrix} \tag{12}$$

where the definition of the primed terms is obvious. Using the symmetry of (12), it will be sufficient to post multiply (11) and (12) by the single vector

$$\begin{Bmatrix} 1 - q^2 \\ tq - s \\ sq - t \\ 0 \end{Bmatrix}$$

to get

$$\begin{bmatrix} 1 + 2\,stq - q^2 - s^2 - t^2 \\ 0 \\ 0 \\ r - 2\,st - rq^2 + t^2q + s^2q \end{bmatrix} \tag{11'}$$

and

$$\begin{bmatrix} (1-q^2)A' + (tq-s)E' + (sq-t)G' \\ (1-q^2)E' + (tq-s)B' + (sq-t)I' \\ (1-q^2)G' + (tq-s)I' + (sq-t)B' \\ (1-q^2)J' + (tq-s)G' + (sq-t)E' \end{bmatrix} \tag{12'}$$

Premultiplying by $\{ 1-q^2 \quad tq-s \quad sq-t \quad 0 \}$, we get

$$(1 + 2\,st\,q - q^2 - s^2 - t^2) (1 - q^2) \tag{11''}$$

and

$$\begin{aligned}&(1-q^2)^2 A' + 2(tq-s)(1-q^2) E' + 2(sq-t)(1-q^2) G' \\ &\quad + (tq-s)^2 B' + 2(sq-t)(tq-s) I' + (sq-t)^2 B'\end{aligned} \tag{12''}$$

Using $2E' = (A'+B')s$, $2G' = (A'+B')t$ and $2I' = 2B' q$, (12") becomes

$$\begin{aligned}&A' [(1-q^2)^2 + s(tq-s)(1-q^2) + (sq-t)(1-q^2)t] + \\ &+ B' [s(tq-s)(1-q^2) + (sq-t)(1-q^2)t + (tq-s)^2 + 2(sq-t)(tq-s)q \\ &\qquad + (sq-t)^2] \\ &= A' (1-q^2) [1 + 2stq - s^2 - t^2 - q^2] \\ &+ B' [(1-q^2)(2stq - s^2 - t^2) + (tq-s)(tq-s + sq^2 - tq) \\ &+ (sq-t)(sq-t + tq^2 - sq)] = A' (1-q^2)[1 + 2stq - s^2 - t^2 - q^2] .\end{aligned}$$

Thus (11") and 12") differ only by the factor A+D .

Premultiplying by $\{ 0 \quad sq-t \quad tq-s \quad 1-q^2 \}$, we get

$$(1-q^2)(r - 2st - rq^2 + t^2q + s^2q) \tag{11'''}$$

and

$$\begin{aligned}&2(sq-t)(1-q^2)E' + 2(tq-s)(sq-t)B' + (sq-t)^2I' + (tq-s)^2 I' \\ &\qquad 2(tq-s)(1-q^2)G' + (1-q^2)^2 J'\end{aligned} \tag{12'''}$$

Dividing (11"') by (11"), our correlation coefficient is given by

$$\frac{r - 2sr - rq^2 + t^2q + s^2q}{1 + 2stq - s^2 - t^2 - q^2} \quad .$$

Equating this to (12"') divided by (12"), we see that we need

$$A'(1-q^2)(\ r-2st - rq^2 + t^2q + s^2q\ ) \ = \ (12''') \quad . \qquad (13)$$

Using $2E' = (A'+B')s$, $I' = B'q$, $2G' = (A'+B')t$, (12"') becomes

$$A'\ [\ (sq-t)(1-q^2)s + (tq-s)(1-q^2)t + (1-q^2)^2\ J'/A'\ ]$$

$$+ B'\ [\ (sq-t)(1-q^2)s + 2(tq-s)(sq-t) + (sq-t)^2q + (tq-s)^2q + (tq-s)(1-q^2)t\ ]$$

$$= A'\ (1-q^2)[\ s^2q - st + t^2q - st + (1-q^2)\ J'/A'\ ]$$

$$+ \ B'[(sq-t)(s-q^2s + tq-s) + (tq-s)(\ sq-t + t - q^2t\ )$$

$$+ (sq-t)^2q + (tq-s)^2\ q\ ]$$

$$= A'\ (1-q^2)\ [\ -2st + s^2q + t^2q + (1-q^2)\ J'/A'\ ]\ . \qquad (12'''')$$

Comparing (13) and (12""), we see that

$$r \ = \ J'/A' \ = \ 2J/(A+D)\ .$$

Again, this is the answer we would have reached if we had completely ignored x_2 and x_3.

Our final transformation of (1) is simply

$$\frac{(A+B+C+D)}{4}\begin{bmatrix} 1 & \frac{2E+2F}{A+B+C+D} & \frac{2G+2H}{A+B+C+D} & \frac{2J}{A+D} \\ \frac{2E+2F}{A+B+C+D} & 1 & \frac{2I}{B+C} & \frac{2G+2H}{A+B+C+D} \\ \frac{2G+2H}{A+B+C+D} & \frac{2I}{B+C} & 1 & \frac{2E+2F}{A+B+C+D} \\ \frac{2J}{A+D} & \frac{2G+2H}{A+B+C+D} & \frac{2E+2F}{A+B+C+D} & 1 \end{bmatrix}.$$

We note that using our general principles of covariance estimation for this special problem are order invarient. I.e., we would get the same answer independent of the order in which we calculated our correlation coefficients.