1. This problem is a particular case of Problem 3 and is addressed there.

2. Consider two measurements of one unknown variable x with correlated noise. Specifically, suppose that

$$y_1 = x + \nu_1,$$
$$y_2 = x + \nu_2,$$

where

$$\nu_1 = \varepsilon_1 + \varepsilon_0,$$
$$\nu_2 = \varepsilon_2 + \varepsilon_0,$$
$$\varepsilon_1, \varepsilon_2 \sim (0, \sigma_1^2), \quad \varepsilon_0 \sim (0, \sigma_0^2),$$
$$\sigma_0^2 + \sigma_1^2 = \sigma^2, \quad r = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.$$

(a) Write it in matrix form

$$y = Ax + \nu$$

and write the matrices A and $S = \mathrm{Var}(\nu)$.

Solution

Let $y = [y_1,\, y_2]^T$, $A = [1,\, 1]^T$ and

$$\nu = \begin{bmatrix} \varepsilon_1 + \varepsilon_0 \\ \varepsilon_2 + \varepsilon_0 \end{bmatrix}.$$

Then, the standard basis matrix representation for the variance operator of the random vector ν is given by

$$S = \begin{bmatrix} \mathbf{E}\,(\varepsilon_1 + \varepsilon_0)^2 & \mathbf{E}\,(\varepsilon_1 + \varepsilon_0)(\varepsilon_2 + \varepsilon_0) \\ \mathbf{E}\,(\varepsilon_1 + \varepsilon_0)(\varepsilon_2 + \varepsilon_0) & \mathbf{E}\,(\varepsilon_1 + \varepsilon_0)^2 \end{bmatrix} = \begin{bmatrix} \sigma_1^2 + \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_1^2 + \sigma_0^2 \end{bmatrix} = \sigma^2 \begin{bmatrix} 1 & r \\ r & 1 \end{bmatrix}.$$

□

(b) Find $\mathrm{Var}(\widehat{x})$ for the optimal linear estimate of x.

Solution

The best linear unbiased estimator is given by

$$R = (A^* S^{-1} A)^{-1} A^* S^{-1}.$$

Observe

$$\begin{aligned}
\mathrm{Var}(\widehat{x}) = \mathrm{Var}(Ry) = \mathrm{Var}(R\nu) &= RSR^* \\
&= \Big((A^*S^{-1}A)^{-1}A^*S^{-1}\Big) S \Big(S^{-1}A(A^*S^{-1}A)^{-1}\Big) \\
&\overset{*}{=} (A^*S^{-1}A)^{-1} \\
&= \left([1\ 1]\left(\frac{1}{\sigma^2(1-r^2)}\begin{bmatrix} 1 & -r \\ -r & 1 \end{bmatrix}\right)\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1} \\
&= (1-r^2)\sigma^2 \left([1-r\ \ 1-r]\begin{bmatrix} 1 \\ 1 \end{bmatrix}\right)^{-1} = \frac{\sigma^2(1+r)}{2}
\end{aligned}$$

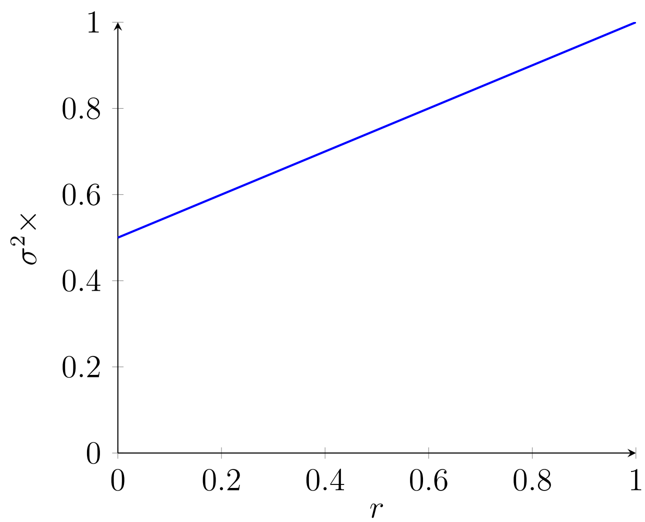
Note that the best linear unbiased estimate is also the least squares estimate,

$$\widehat{x} = Ry = \mathrm{Var}(\widehat{x})A^* S^{-1} y = \left(\frac{\sigma^2/2}{1-r}\right)\left(\frac{1}{\sigma^2}[1-r \;\; 1-r]\right) y = \frac{y_1+y_2}{2}.$$

$\square$

(c) Analyze how the variance of $\widehat{x}$ depends on the correlation parameter r for $0 \leq r \leq 1$. Is higher correlation good or bad for estimation in this example? A graph might be helpful. How would you explain such behavior?

Below, we plot the variance of $\widehat{x}$ as a function of r.



In this scheme, measurements are taken independently, and, intuitively, one would expect correlation between measurements to reduce the information in the two measurements. Indeed, the plot above shows that as $r \to 1$, the uncertainty in the estimate increases until the case when $\sigma_0 \gg \sigma_1$, in which case the "slow" error term dominates the "fast" one so that it is as if only one replication has occurred, i.e. $\nu_1 \approx \nu_2$, and we have only one measurement of the random quantity $x + \nu_1$, and our estimate has variance σ^2.

3. Consider the following set of measurements of the unknown variables x_1 and x_2:

$$\begin{aligned} y_1 &= x_1 + x_2 + \nu_1 \\ y_2 &= x_1 - x_2 + \nu_2 \\ y_3 &= -x_1 + x_2 + \nu_3 \end{aligned}$$

where y_i are measurement results, and

$$\begin{aligned} \nu_1 &= \varepsilon_1 + \varepsilon_0, \\ \nu_2 &= \varepsilon_2 + \varepsilon_0, \\ \nu_3 &= \varepsilon_3 + \varepsilon_0, \end{aligned}$$

$$\varepsilon_1, \varepsilon_2, \varepsilon_3 \sim (0, \sigma_1^2), \quad \varepsilon_0 \sim (0, \sigma_0^2),$$
$$\sigma_0^2 + \sigma_1^2 = \sigma^2, \quad r = \frac{\sigma_0^2}{\sigma_0^2 + \sigma_1^2}.$$

(a) Write it in matrix form

$$y = Ax + \nu$$

and write the matrices A and $S = \mathrm{Var}(\nu)$.

Solution

Let $y = [y_1\ y_2\ y_3]^T, x = [x_1\ x_2]^T$,

$$\nu = \begin{bmatrix} \varepsilon_1 + \varepsilon_0 \\ \varepsilon_2 + \varepsilon_0 \\ \varepsilon_3 + \varepsilon_0 \end{bmatrix}, \quad \text{and} \quad A = \begin{bmatrix} 1 & 1 \\ 1 & -1 \\ -1 & 1 \end{bmatrix}.$$

The standard basis representation of the variance operator of ν is

$$\begin{aligned} S &= \begin{bmatrix} \mathbf{E}\,(\varepsilon_1+\varepsilon_0)^2 & \mathbf{E}\,(\varepsilon_1+\varepsilon_0)(\varepsilon_2+\varepsilon_0) & \mathbf{E}\,(\varepsilon_1+\varepsilon_0)(\varepsilon_3+\varepsilon_0) \\ \mathbf{E}\,(\varepsilon_1+\varepsilon_0)(\varepsilon_2+\varepsilon_0) & \mathbf{E}\,(\varepsilon_1+\varepsilon_0)^2 & \mathbf{E}\,(\varepsilon_1+\varepsilon_0)(\varepsilon_3+\varepsilon_0) \\ \mathbf{E}\,(\varepsilon_1+\varepsilon_0)(\varepsilon_3+\varepsilon_0) & \mathbf{E}\,(\varepsilon_2+\varepsilon_0)(\varepsilon_3+\varepsilon_0) & \mathbf{E}\,(\varepsilon_3+\varepsilon_0)^2 \end{bmatrix} \\ &= \begin{bmatrix} \sigma_1^2+\sigma_0^2 & \sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_1^2+\sigma_0^2 & \sigma_0^2 \\ \sigma_0^2 & \sigma_0^2 & \sigma_1^2+\sigma_0^2 \end{bmatrix} \\ &= \sigma^2 \begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix}. \end{aligned}$$

In the case where measurement error is independent, this corresponds fixing $\varepsilon_0 \equiv 0$ which implies that $r = 0$, so $S = \sigma^2 I$.

□

(b) Find the variance matrix $\mathrm{Var}(\widehat{x})$ for the optimal linear estimate of x and variances $\widehat{x}_1$ and $\widehat{x}_2$.

Solution

(Remark: These computations were done with the assistance of the computer algebra system `Maxima`). Let us first calculate

$$A^* S^{-1} = \frac{1}{\sigma^2} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix}^{-1}.$$

Observe,

$$\begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix} \begin{bmatrix} -(1+r) & r & r \\ r & -(1+r) & r \\ r & r & -(1+r) \end{bmatrix} = \begin{bmatrix} 2r^2-r-1 & 0 & 0 \\ 0 & 2r^2-r-1 & 0 \\ 0 & 0 & 2r^2-r-1 \end{bmatrix}$$

so

$$\begin{aligned} A^* S^{-1} &= \frac{1}{\sigma^2} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & r & r \\ r & 1 & r \\ r & r & 1 \end{bmatrix}^{-1} \\ &= \frac{1}{\sigma^2(2r-r-1)} \begin{bmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \end{bmatrix} \begin{bmatrix} -(1+r) & r & r \\ r & -(1+r) & r \\ r & r & -(1+r) \end{bmatrix} \\ &= \frac{1}{\sigma^2(2r+1)(r-1)} \begin{bmatrix} -(r+1) & -(r+1) & 3r+1 \\ -(r+1) & 3r+1 & -(r+1) \end{bmatrix}. \end{aligned}$$

Thus,

$$
\begin{aligned}
\operatorname{Var}(\widehat{x}) &\overset{*}{=} (A^* S^{-1} A)^{-1} \\
&= \sigma^2 (2r+1)(r-1) \left(\begin{bmatrix} -5r-3 & 3r+1 \\ 3r+1 & -5r-3 \end{bmatrix} \right)^{-1} \\
&= \frac{\sigma^2(2r+1)(r-1)}{(5r+3)^2 - (3r+1)^2} \begin{bmatrix} -5r-3 & -3r-1 \\ -3r-1 & -5r-3 \end{bmatrix} \\
&= \frac{\sigma^2(2r+1)(1-r)}{8(r+1)(2r+1)} \begin{bmatrix} 5r+3 & 3r+1 \\ 3r+1 & 5r+3 \end{bmatrix} \\
&= \frac{\sigma^2(1-r)}{8(r+1)} \begin{bmatrix} 5r+3 & 3r+1 \\ 3r+1 & 5r+3 \end{bmatrix}.
\end{aligned}
$$

Note that the symmetric measurement scheme leads to identical variance for the optimal estimates,

$$\operatorname{Var}(\widehat{x}_1) = \operatorname{Var}(\widehat{x}_2) = \sigma^2 \frac{(1-r)(5r+3)}{8(r+1)}.$$

The optimal estimates are given by

$$
\begin{aligned}
\widehat{x} &= Ry \\
&= \operatorname{Var}(\widehat{x}) A^* S^{-1} y \\
&= \left(\frac{\sigma^2(1-r)}{8(r+1)} \begin{bmatrix} 5r+3 & 3r+1 \\ 3r+1 & 5r+3 \end{bmatrix} \right) \left(\frac{1}{\sigma^2(2r+1)(r-1)} \begin{bmatrix} -(r+1) & -(r+1) & 3r+1 \\ -(r+1) & 3r+1 & -(r+1) \end{bmatrix} \right) y \\
&= \frac{1}{8(2r+1)(r+1)} \begin{bmatrix} 5r+3 & 3r+1 \\ 3r+1 & 5r+3 \end{bmatrix} \begin{bmatrix} r+1 & r+1 & -(3r+1) \\ r+1 & -(3r+1) & r+1 \end{bmatrix} y \\
&= \frac{1}{8(2r+1)(r+1)} \begin{bmatrix} 8r^2+12r+4 & -4r^2+2r+2 & -12r^2-10r-2 \\ 8r^2+12r+4 & -12r^2-10r-2 & -4r^2+2r+2 \end{bmatrix} y \\
&= \frac{1}{8(2r+1)(r+1)} \begin{bmatrix} 4(r+1)(2r+1) & -2(r-1)(2r+1) & -2(2r+1)(3r+1) \\ 4(r+1)(2r+1) & -2(2r+1)(3r+1) & -2(r-1)(2r+1) \end{bmatrix} y \\
&= \frac{1}{4} \begin{bmatrix} 2 & -\dfrac{r-1}{r+1} & -\dfrac{3r+1}{r+1} \\ 2 & -\dfrac{3r+1}{r+1} & -\dfrac{r-1}{r+1} \end{bmatrix} y
\end{aligned}
$$

Note that these estimates depend on r, so if it is unknown, further work is required to obtain an estimate $\widehat{r}$.

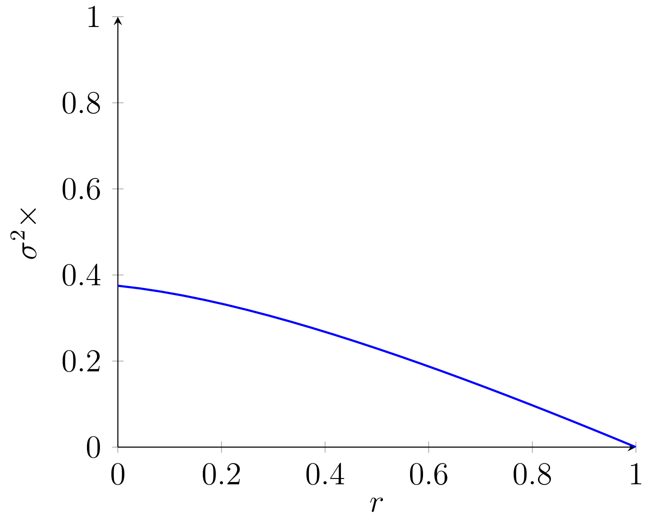
These reduce to the following expressions when the noise is assumed to be independent:

$$\widehat{x} = \frac{1}{4} \begin{bmatrix} 2 & 1 & -1 \\ 2 & -1 & 1 \end{bmatrix} y, \ \operatorname{Var}(\widehat{x}) = \frac{\sigma^2}{8} \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \quad \text{and} \quad \operatorname{Var}(\widehat{x}_1) = \operatorname{Var}(\widehat{x}_2) = \frac{3}{8}\sigma^2.$$

□

(c) Analyze how the variances of $\widehat{x}_1$ and $\widehat{x}_2$ depend on the correlation parameter r for $0 \leq r \leq 1$. Is higher correlation good or bad for estimation in this example? A graph might be helpful. How would you explain such behavior?

Below, we plot the variances of the estimates for $\widehat{x}_1$ and $\widehat{x}_2$.



Note that in this measurement scheme, higher correlation is *better* for estimating. Each measurement depends on the others, since each indirectly involves both x_1 and x_2, and, in some sense, this measurement scheme "balances" the error relationship in a way that gives more information when there is higher correlation between measurements. In the limit as $\sigma_0 \gg \sigma_1$, we have that given the three measurements y_1, y_2, and y_3, the three quantities x_1, x_2, and $\nu_1 = \nu_2 = \nu_3$ are known exactly and there is no uncertainty in the estimates.