Definition 1. *Let X and Y be two column vectors of length n. Then the sample covariance between X and Y, $Cov[X,Y]$, is defined to be*

$$Cov[X,Y] = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \mu_X)(y_i - \mu_Y),$$

where μ_X is the mean of X and μ_Y is the mean of Y.

Definition 2. *Let*

$$X = \begin{bmatrix} X_1 & X_2 & \cdots & X_m \end{bmatrix}^T,$$

where X_i is a column vector. Then, the covariance matrix of X is

$$\begin{bmatrix} Cov[X_1,X_1] & Cov[X_1,X_2] & \cdots & Cov[X_1,X_m] \\ Cov[X_2,X_1] & Cov[X_2,X_2] & \cdots & Cov[X_2,X_m] \\ \vdots & \vdots & \ddots & \vdots \\ Cov[X_m,X_1] & Cov[X_m,X_2] & \cdots & Cov[X_m,X_m] \end{bmatrix}$$

Lemma 1. *Let X be defined as above. Let μ be the column vector of the means of each $X_i \in X$.*

$$\mu = \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix} = \begin{bmatrix} \frac{1}{n}\sum_{i=1}^{n} x_{1i} \\ \frac{1}{n}\sum_{i=1}^{n} x_{2i} \\ \vdots \\ \frac{1}{n}\sum_{i=1}^{n} x_{mi} \end{bmatrix}.$$

Then, the covariance matrix of X may be written as

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T.$$

Proof.

$$\frac{1}{n-1}\sum_{i=1}^{n}(X_i - \mu)(X_i - \mu)^T = \frac{1}{n-1}\sum_{i=1}^{n}\begin{bmatrix} x_{1i}-\mu_1 \\ x_{2i}-\mu_2 \\ \vdots \\ x_{mi}-\mu_m \end{bmatrix}\begin{bmatrix} x_{1i}-\mu_1 & x_{2i}-\mu_2 & \cdots & x_{mi}-\mu_m \end{bmatrix}$$

$$= \frac{1}{n-1}\sum_{i=1}^{n}\begin{bmatrix} (x_{1i}-\mu_1)(x_{1i}-\mu_1) & (x_{1i}-\mu_1)(x_{2i}-\mu_2) & \cdots & (x_{1i}-\mu_1)(x_{mi}-\mu_m) \\ (x_{2i}-\mu_1)(x_{1i}-\mu_1) & (x_{2i}-\mu_1)(x_{2i}-\mu_2) & \cdots & (x_{2i}-\mu_1)(x_{mi}-\mu_m) \\ \vdots & \vdots & \ddots & \vdots \\ (x_{mi}-\mu_1)(x_{1i}-\mu_1) & (x_{mi}-\mu_1)(x_{2i}-\mu_2) & \cdots & (x_{mi}-\mu_1)(x_{mi}-\mu_m) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i}-\mu_1)(x_{1i}-\mu_1) & \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i}-\mu_1)(x_{2i}-\mu_2) & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{1i}-\mu_1)(x_{mi}-\mu_m) \\ \frac{1}{n-1}\sum_{i=1}^{n}(x_{2i}-\mu_1)(x_{1i}-\mu_1) & \frac{1}{n-1}\sum_{i=1}^{n}(x_{2i}-\mu_1)(x_{2i}-\mu_2) & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{2i}-\mu_1)(x_{mi}-\mu_m) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{1}{n-1}\sum_{i=1}^{n}(x_{mi}-\mu_1)(x_{1i}-\mu_1) & \frac{1}{n-1}\sum_{i=1}^{n}(x_{mi}-\mu_1)(x_{2i}-\mu_2) & \cdots & \frac{1}{n-1}\sum_{i=1}^{n}(x_{mi}-\mu_1)(x_{mi}-\mu_m) \end{bmatrix}$$

which is the covariance matrix of X. □

Lemma 2.

$$\sum_{i=1}^{n}(X_i-\mu)=0.$$

Proof.

$$\begin{aligned} \sum(X_i-\mu) &= \sum_{i=1}^{n}\left(\begin{bmatrix} x_{1i} \\ x_{2i} \\ \vdots \\ x_{mi} \end{bmatrix} - \begin{bmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_m \end{bmatrix}\right) \\ &= \sum_{i=1}^{n}\begin{bmatrix} x_{1i}-\mu_1 \\ x_{2i}-\mu_2 \\ \vdots \\ x_{mi}-\mu_m \end{bmatrix} \\ &= \begin{bmatrix} \sum_{i=1}^{n}(x_{1i}-\mu_1) \\ \sum_{i=1}^{n}(x_{2i}-\mu_2) \\ \vdots \\ \sum_{i=1}^{n}(x_{mi}-\mu_m) \end{bmatrix} \\ &= \begin{bmatrix} n\mu_1-n\mu_1 \\ n\mu_2-n\mu_2 \\ \vdots \\ n\mu_m-n\mu_m \end{bmatrix} \\ &= 0 \end{aligned}$$

□

Corollary 1.

$$\sum_{i=1}^{n}(X_i-\mu)^T=0.$$

by identical proof construction as Lemma 1.

Theorem 1. *Let $F(X)$ be the covariance matrix of X without the normalization factor of $n-1$. Let M_i be a maximum bound on $x_i \in X_i$, and let m_i be a minimum bound on $x_i \in X_i$. Then each entry f_{ij} of this matrix has sensitivity bounded above by*

$$\frac{2(n-1)}{n}(M_i - m_i)(M_j - m_j)$$

Proof. Let X be defined as above and let X' be defined as

$$X' = \begin{bmatrix} X_1' & \cdots & X_m' \end{bmatrix}^T$$

where

$$X_i' = X_i \cup \{y_i\}.$$

I.e., each row i has a single additional observation y_i in X' that it does not have in X. Let Y be the vector of all these additional observations. Then,

$$F(X') = \sum (X_i - \mu')(X_i - \mu')^T + (Y - \mu')(Y - \mu')^T.$$

The first of the sums inside this expression may be expanded to give

$$\begin{aligned}
\sum (x_i - \mu')(x_i - \mu')^T &= \sum ((x_i - \mu) + (\mu - \mu'))((x_i - \mu) + (\mu - \mu'))^T \\
&= \sum (x_i - \mu)(x_i - \mu)^T + (\mu - \mu') \sum (x_i - \mu)^T + \sum (x_i - \mu)(\mu - \mu')^T \\
&\quad + \sum (\mu - \mu')(\mu - \mu')^T \\
&= \sum (x_i - \mu)(x_i - \mu)^T + (\mu - \mu') \sum (x_i - \mu)^T + \sum (x_i - \mu)(\mu - \mu')^T \\
&\quad + n(\mu - \mu')(\mu - \mu')^T \\
&= \sum (x_i - \mu)(x_i - \mu)^T + n(\mu - \mu')(\mu - \mu')^T \\
&= F(X) + n(\mu - \mu')(\mu - \mu')^T,
\end{aligned}$$

where the second-to-last line is due to cancellations of the middle two terms by Lemma 2 and Corollary 1. So,

$$F(X') = F(X) + n(\mu - \mu')(\mu - \mu')^T + (Y - \mu')(Y - \mu')^T. \tag{1}$$

Looking at the two expressions inside the parentheses of Eq. 1, note first that

$$n(\mu - \mu')(\mu - \mu')^T$$

is an $m \times m$ matrix with ijth entry

$$\begin{aligned}
x_{ij} &= n(\mu_i - \mu_i')(\mu_j - \mu_j') \\
&\leq n \left(\frac{M_i - m_i}{n+1} \right) \left(\frac{M_j - m_j}{n+1} \right) \\
&= \frac{n}{(n+1)^2}(M_i - m_i)(M_j - m_j).
\end{aligned} \tag{2}$$

The second term,

$$(Y - \mu')(Y - \mu')^T,$$

is also an $m \times m$ matrix, with ijth entry

$$\begin{aligned} x_{ij} &= (y_i - \mu_i')(y_j - \mu_j') \\ &= \left(y_i - \frac{n\mu_i + y_i}{n+1}\right)\left(y_j - \frac{n\mu_j + y_j}{n+1}\right) \\ &= \frac{n^2}{(n+1)^2}(y_i - \mu_i)(y_j - \mu_j) \\ &\leq \frac{n^2}{(n+1)^2}(M_i - m_i)(M_j - m_j). \end{aligned} \tag{3}$$

Let f_{ij} be the ijth entry of the $m \times m$ matrix output by F. Then plugging the bounds in Eq. 2 and Eq. 3 back into Eq. 1 gives

$$\begin{aligned} f_{ij}(X') &\leq f_{ij}(X) + \frac{n}{(n+1)^2}(M_i - m_i)(M_j - m_j) + \frac{n^2}{(n+1)^2}(M_i - m_i)(M_j - m_j) \\ &= f_{ij}(X) + \frac{n}{(n+1)^2}(M_i - m_i)(M_j - m_j)(n+1) \\ &= f_{ij}(X) + \frac{n}{n+1}(M_i - m_i)(M_j - m_j). \end{aligned} \tag{4}$$

Since we'd really like to consider the sensitivity of X', it makes sense to redefine n based on the size of X' rather than of X, i.e. redefine n to be $n+1$. Then,

$$f_{ij}(X') = f_{ij}(X) + \frac{n-1}{n}(M_i - m_i)(M_j - m_j). \tag{5}$$

Now, consider two neighboring databases X' and X''. Say X' may still be written as $X \cup \{y\}$, and X'' may be similarly written as $X \cup \{z\}$. It then follows from Eq. 5, using the triangle inequality, that

$$\left|f_{ij}(X') - f_{ij}(X'')\right| \leq \frac{2(n-1)}{n}(M_i - m_i)(M_j - m_j).$$

Can get tighter maybe? (Get rid of the 2?) Try redoing analysis of Eq. 2 with y and z maybe?

$\square$

Theorem 2. *Consider the case in the PSI-library where a user wants to compute a covariance matrix including the intercept. Then, if X is the original matrix of data values input by the user, the covariance will be calculated over*

$$Y = \begin{bmatrix} \mathbb{1} & X \end{bmatrix}.$$

Then, all $(i, 1)$ and $(1, i)$ entries have sensitivity 0.

Proof. Let Y and Y' be neighboring databases. Note that Y' will not differ in the first column of 1s, as this column is added to all input databases. Then, for an arbitrary column i of X, $\mathrm{Cov}[\mathbb{1}, X_i] = 0$, and $\mathrm{Cov}[\mathbb{1}, X_i'] = 0$, so the sensitivity is 0 for any covariance over the first column of Y. □