Identities for Sums of Squares

1. SSXX $= \sum_{i=1}^{n} (x_i - \overline{x})^2$ [definition]

2. SSXX $= \sum_{i=1}^{n} (x_i - \overline{x})x_i$

3. SSXX $= \sum_{i=1}^{n} x_i^2 - n\overline{x}^2$

4. SSXX $= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\sum_{i=1}^{n} x_i)^2$

5. SSXX $= \frac{1}{n} \cdot \sum_{1 \le j < k \le n} (x_j - x_k)^2$

6. SSXX $= -2 \sum_{j<k} (x_j - \overline{x})(x_k - \overline{x})$

7. SSXX $= \frac{1}{n}\det\left(X^T X\right)$ where $X^T = \begin{bmatrix} 1 & 1 & \dots & 1 \\ X_1 & X_2 & \dots & X_n \end{bmatrix}$

8. SSXY $= \sum_{i=1}^{n} (x_i - \overline{x})(y_i - \overline{y})$ [definition]

9. SSXY $= \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{j=i}^{n} y_j / n$

10. SSXY $= \frac{1}{n} \sum_{j<k} (x_j - x_k)(y_j - y_k)$

11. $\frac{\text{SSXY}}{\text{SSXX}} = \frac{\sum_{j<k}(x_j - x_k)(y_j - y_k)}{\sum_{j<k}(x_j - x_k)(x_j - x_k)}$

12. SSYY $= \sum_{i=1}^{n} (y_i - \overline{y})^2$ [definition]

13. SSYY $= \boldsymbol{Y}^T(\boldsymbol{I} - \frac{1}{n}\boldsymbol{J})\boldsymbol{Y}$ where $\boldsymbol{Y}^T = [Y_1\ Y_2\ \dots Y_n]$, $\boldsymbol{I}$ is the $n \times n$ identity matrix, and $\boldsymbol{J}$ is the $n \times n$ matrix consisting of all ones.

Proof of 2.

$$\begin{aligned}
\text{SSXX} &= \sum_{i=1}^{n} (x_i - \overline{x})^2 \\
&= \sum_{i=1}^{n} (x_i - \overline{x})x_i - \overline{x}\sum_{i=1}^{n}(x_i - \overline{x}) \\
&= \sum_{i=1}^{n} (x_i - \overline{x})x_i \quad \text{[since } \sum_{i=1}^{n}(x_i - \overline{x}) = 0] \qquad \square
\end{aligned}$$

Proof of 3, 4.

$$\begin{aligned}
\text{SSXX} &= \sum_{i=1}^{n} (x_i - \overline{x})x_i \\
&= \sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i \\
&= \sum_{i=1}^{n} x_i^2 - n\,\overline{x}^2 \qquad (3) \\
&= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\,(\sum_{i=1}^{n} x_i)^2 \qquad (4) \qquad \square
\end{aligned}$$

Proof 5.

$$\begin{aligned}
\frac{1}{n}\cdot\sum_{j<k} (x_j - x_k)^2 &= \frac{1}{n}\cdot\sum_{j<k}(x_j^2 + x_k^2) - \frac{1}{n}\sum_{j<k} 2x_j x_k \\
&= \frac{1}{n}\,(n-1)\sum_{i=1}^{n} x_i^2 - \frac{1}{n}(\,(\sum_{i=1}^{n} x_i)^2 - \sum_{i=1}^{n} x_i^2) \\
&= \sum_{i=1}^{n} x_i^2 - \frac{1}{n}\,(\sum_{i=1}^{n} x_i)^2 \\
&= \text{SSXX} \qquad \square
\end{aligned}$$

Proof 6.

$$
\begin{aligned}
-2\sum_{i<j}(x_i-\overline{x})(x_j-\overline{x}) &= 2\overline{x}\sum_{i<j}(x_i+x_j)-2\sum_{i<j}x_ix_j-2\binom{n}{2}\overline{x}^{\,2}\\
&= 2\overline{x}\,(n-1)\sum_{i=1}^{n}x_i-\left(\left(\sum x_i\right)^2-\sum x_i^2\right)-n(n-1)\,\overline{x}^{\,2}\\
&= 2\overline{x}^{\,2}n(n-1)-n^2\overline{x}^2+\sum x_i^2-n(n-1)\,\overline{x}^{\,2}\\
&= \overline{x}^{\,2}\big(2n(n-1)-n^2-n(n-1)\big)+\sum x_i^2\\
&= -n\,\overline{x}^{\,2}+\sum x_i^2\\
&= \text{SSXX} \qquad \square
\end{aligned}
$$

Proof of 7.

$$
\begin{aligned}
\tfrac{1}{n}\det\left(X^TX\right) &= \frac{1}{n}\begin{vmatrix} n & \sum x_i \\ \sum x_i & \sum x_i^2 \end{vmatrix}\\
&= \tfrac{1}{n}\left(n\sum x_i^2-\left(\sum x_i\right)^2\right.\\
&= \text{SSXX} \qquad \square
\end{aligned}
$$

Proof of 9.

$$
\begin{aligned}
\text{SSXY} &= \sum_{i=1}^{n}(x_i-\overline{x})(y_i-\overline{y})\\
&= \sum_{i=1}^{n}(x_i-\overline{x})y_i \qquad \left[\text{since } \overline{y}\sum_{i=1}^{n}(x_i-\overline{x})=0\right]\\
&= \sum_{i=1}^{n}x_iy_i-\overline{x}\sum_{i=1}^{n}y_i\\
&= \sum_{i=1}^{n}x_iy_i-\tfrac{1}{n}\sum_{i=1}^{n}x_i\sum_{i=1}^{n}y_i \qquad \square
\end{aligned}
$$

Proof of 10.

$$\frac{1}{n}\sum_{j<k}(x_j - x_k)(y_j - y_k) = \frac{1}{n}\sum_{j<k}(x_j y_j - x_k y_j - x_j y_k + x_k y_k)$$

$$= \frac{1}{n}\sum_{k=2}^{n}\sum_{j=1}^{k-1}(x_j y_j + x_k y_k) - \frac{1}{n}\sum_{k=2}^{n}\sum_{j=1}^{k-1}(x_k y_j + x_j y_k)$$

$$= \frac{n-1}{n}\sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{k=2}^{n}\sum_{j=1}^{k-1}(x_k y_j + x_j y_k)$$

$$= \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\Big(\sum_{i=1}^{n} x_i y_i + \sum_{k=2}^{n}\sum_{j=1}^{k-1}(x_k y_j + x_j y_k)\Big)$$

$$= \sum_{i=1}^{n} x_i y_i - \frac{1}{n}\sum_{i=1}^{n} x_i \sum_{j=1}^{n} y_i$$

$$= \text{SSXY} \qquad \square$$

Proof of 11. $\frac{\text{SSXY}}{\text{SSXX}} = \frac{\frac{1}{n}\sum_{j<k}(x_j - x_k)(y_j - y_k)}{\frac{1}{n}\sum_{j<k}(x_j - x_k)^2}$ [from **5** and **10**]

$$= \frac{\sum_{j<k}(x_j - x_k)(y_j - y_k)}{\sum_{j<k}(x_j - x_k)(x_j - x_k)} \qquad \square$$