Exercise: FAST Interaction Detection with Prefix Sums

Exercise

The table lists $n = 9$ samples with two numerical features and a target:

Idx	x_1	x_2	y
0	1.07	1.11	2
1	1.86	1.05	3
2	3.18	1.07	4
3	0.93	2.16	5
4	2.12	2.08	6
5	2.85	2.14	7
6	1.18	3.06	8
7	2.03	2.92	9
8	3.09	3.17	10

Tasks

1. Discretize each feature into three equal-width bins $[0, 1.5) \to 0,\ [1.5, 2.5) \to 1,\ [2.5, 3.5) \to 2$.
2. Build two 3×3 matrices

$$S(i,j) = \sum_{(x_1^b, x_2^b)=(i,j)} y, \qquad N(i,j) = \sum_{(x_1^b, x_2^b)=(i,j)} 1.$$

3. Form 2-D prefix sums $S^{\text{pref}}, N^{\text{pref}}$.
4. Via inclusion–exclusion obtain totals for the rectangle $x_1^b \in \{1,2\}$, $x_2^b \in \{1,2\}$: $S_r,\ N_r$.
5. Compute the mean in this region: $\hat{y}_r = S_r / N_r$.
6. Let $S_n = \sum_{i=1}^{n} y^{(i)},\ N_n = n$. Show that the *RSS reduction*

$$\Delta\text{RSS} = \frac{S_r^2}{N_r} + \frac{(S_n - S_r)^2}{N_n - N_r} - \frac{S_n^2}{N_n}$$

   needs only first-order sums (no y^2).

Solution

1. Binning the data

Idx	x_1	x_2	y	(x_1^b, x_2^b)
0	1.07	1.11	2	(0,0)
1	1.86	1.05	3	(1,0)
2	3.18	1.07	4	(2,0)
3	0.93	2.16	5	(0,1)
4	2.12	2.08	6	(1,1)
5	2.85	2.14	7	(2,1)
6	1.18	3.06	8	(0,2)
7	2.03	2.92	9	(1,2)
8	3.09	3.17	10	(2,2)

2. Aggregate matrices S and N

$$S = \begin{bmatrix} 2 & 5 & 8 \\ 3 & 6 & 9 \\ 4 & 7 & 10 \end{bmatrix}, \qquad N = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

3. 2-D prefix sums

$$S^{\text{pref}} = \begin{bmatrix} 2 & 7 & 15 \\ 5 & 16 & 33 \\ 9 & 27 & 54 \end{bmatrix}, \qquad N^{\text{pref}} = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 4 & 6 \\ 3 & 6 & 9 \end{bmatrix}$$

4. Rectangle $(1{:}2, 1{:}2)$

Let the binned grid have indices $0, \dots, B-1$ on each axis and define an axis-aligned rectangle

$$R = [i_{\min} : i_{\max}] \times [j_{\min} : j_{\max}], \qquad 0 \le i_{\min} \le i_{\max} < B,\ 0 \le j_{\min} \le j_{\max} < B.$$

With 2-D prefix sums

$$S^{\text{pref}}(i,j) = \sum_{u \le i,\ v \le j} S(u,v), \qquad N^{\text{pref}}(i,j) = \sum_{u \le i,\ v \le j} N(u,v),$$

the totals in R are obtained via inclusion–exclusion:

$$\boxed{\begin{aligned} S_R &= S^{\text{pref}}_{i_{\max}, j_{\max}} - S^{\text{pref}}_{i_{\min}-1, j_{\max}} - S^{\text{pref}}_{i_{\max}, j_{\min}-1} + S^{\text{pref}}_{i_{\min}-1, j_{\min}-1}, \\ N_R &= N^{\text{pref}}_{i_{\max}, j_{\max}} - N^{\text{pref}}_{i_{\min}-1, j_{\max}} - N^{\text{pref}}_{i_{\max}, j_{\min}-1} + N^{\text{pref}}_{i_{\min}-1, j_{\min}-1}. \end{aligned}} \tag{1}$$

(All terms with index -1 are interpreted as zero.)

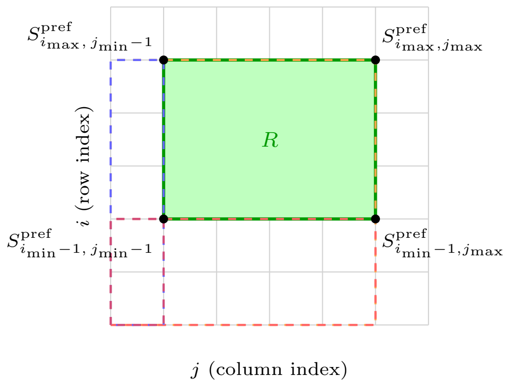

Figure 1: Inclusion–exclusion on a 2-D prefix grid. The green area is the query rectangle R. Dashed rectangles show the four prefix sums employed in Eq. (1): orange (+), red and blue (−), purple (+).

Using the same four look-ups one can obtain the subtotals of the *other three* regions that partition the orange prefix rectangle:

$$\begin{aligned} S_{\text{top}} &= S^{\text{pref}}_{i_{\min}-1, j_{\max}} - S^{\text{pref}}_{i_{\min}-1, j_{\min}-1}, \\ S_{\text{left}} &= S^{\text{pref}}_{i_{\max}, j_{\min}-1} - S^{\text{pref}}_{i_{\min}-1, j_{\min}-1}, \\ S_{\text{topleft}} &= S^{\text{pref}}_{i_{\min}-1, j_{\min}-1}. \end{aligned}$$

(Replace S^{pref} by N^{pref} to get the corresponding counts.)

RSS drop using only first-order sums. Let $S_n = \sum_{i=1}^{n} y^{(i)}$ and $N_n = n$. Define $S_C = S_n - S_R$ and $N_C = N_n - N_R$. Then the reduction in residual sum of squares when isolating rectangle R is

$$\boxed{\Delta \text{RSS}(R) = \frac{S_R^2}{N_R} + \frac{S_C^2}{N_C} - \frac{S_n^2}{N_n}}, \tag{2}$$

which involves *only* the first-order target sums $S_\bullet$ and counts $N_\bullet$; all $\sum y^2$ terms cancel. This identity is the core of the FAST interaction-search algorithm and the 1-D prefix-sum split optimisation presented in the lecture.

Using inclusion–exclusion:

$$S_r = S_{2,2}^{\text{pref}} - S_{0,2}^{\text{pref}} - S_{2,0}^{\text{pref}} + S_{0,0}^{\text{pref}} = 54 - 15 - 9 + 2 = 32,$$
$$N_r = N_{2,2}^{\text{pref}} - N_{0,2}^{\text{pref}} - N_{2,0}^{\text{pref}} + N_{0,0}^{\text{pref}} = 9 - 3 - 3 + 1 = 4.$$

5. Mean prediction

$$\hat{y}_r = \frac{S_r}{N_r} = \frac{32}{4} = 8.$$

6. RSS reduction with first-order sums only

Total sums: $S_n = 54,\ N_n = 9$. Complement: $S_c = S_n - S_r = 22,\ N_c = N_n - N_r = 5$.

$$\boxed{\Delta\text{RSS} = \frac{S_r^2}{N_r} + \frac{S_c^2}{N_c} - \frac{S_n^2}{N_n} = \frac{32^2}{4} + \frac{22^2}{5} - \frac{54^2}{9} = 28.8}$$

Observation: The formula contains only the sums S and counts N; all $\sum y^2$ terms cancel. This is exactly the trick exploited by FAST and the prefix-sum split algorithm in the lecture.