

Question 2-1

guilhermecorreagv

August 2021

Ex 2.1: Least squares intersection point and line fitting—advanced. Equation (2.4) shows how the intersection of two 2D lines can be expressed as their cross product, assuming the lines are expressed as homogeneous coordinates.

1. If you are given more than two lines and want to find a point \tilde{x} that minimizes the sum of squared distances to each line,

$$D = \sum_i (\tilde{x} \cdot \tilde{l}_i)^2$$

how can you compute this quantity? (Hint: Write the dot product as $\tilde{x}^T \tilde{l}_i$ and turn the squared quantity into a quadratic form, $\tilde{x}^T A \tilde{x}$.)

2. To fit a line to a bunch of points, you can compute the centroid (mean) of the points as well as the covariance matrix of the points around this mean. Show that the line passing through the centroid along the major axis of the covariance ellipsoid (largest eigenvector) minimizes the sum of squared distances to the points.
3. These two approaches are fundamentally different, even though projective duality tells us that points and lines are interchangeable. Why are these two algorithms so apparently different? Are they actually minimizing different objectives?

I. SOLUTION

Following the hint given by the author we have:

$$\begin{aligned}
D &= \sum_i \tilde{\mathbf{x}}^T \tilde{\mathbf{l}}_i \tilde{\mathbf{x}}^T \tilde{\mathbf{l}}_i \\
D &= \sum_i \tilde{\mathbf{x}}^T \tilde{\mathbf{l}}_i \left(\tilde{\mathbf{l}}_i^T \tilde{\mathbf{x}} \right)^T \\
D &= \sum_i \tilde{\mathbf{x}}^T \tilde{\mathbf{l}}_i \tilde{\mathbf{l}}_i^T \tilde{\mathbf{x}} \\
D &= \sum_i \tilde{\mathbf{x}}^T A_i \tilde{\mathbf{x}} \\
D &= \tilde{\mathbf{x}}^T \left(\sum_i A_i \right) \tilde{\mathbf{x}} \\
D &= \tilde{\mathbf{x}}^T A \tilde{\mathbf{x}}
\end{aligned}$$

Where

$$A = \sum_i \tilde{\mathbf{l}}_i \tilde{\mathbf{l}}_i^T$$

To avoid the trivial solution we'll restrict ourselves to the case where $\|\mathbf{x}\|_2 = 1$. We also note that by our construction A is semi-positive definite and symmetric. We then build the following Lagrangian:

$$\mathcal{L}(\tilde{\mathbf{x}}, \lambda) = \tilde{\mathbf{x}}^T A \tilde{\mathbf{x}} + \lambda(\mathbf{1} - \tilde{\mathbf{x}}^T \tilde{\mathbf{x}})$$

We have then:

$$\frac{\partial \mathcal{L}(\tilde{\mathbf{x}}, \lambda)}{\partial \tilde{\mathbf{x}}} = 2A\tilde{\mathbf{x}} - 2\lambda\tilde{\mathbf{x}} \quad (A - \lambda I)\tilde{\mathbf{x}} = 0$$

We notice that $\tilde{\mathbf{x}}$ is an eigenvector of A , and thus we have:

$$\tilde{\mathbf{x}}^T A \tilde{\mathbf{x}} = \tilde{\mathbf{x}}^T \lambda^* \tilde{\mathbf{x}} = \lambda^*$$

So the solution is the eigenvector associated with the smallest eigenvalue of A .

We have a matrix X where each row is a point. Then the centroid $\bar{\mathbf{x}}$ is:

$$\bar{\mathbf{x}} = \frac{1}{n} \sum_i X_i$$

We then extend this average row by doing:

$$\bar{X} = \mathbf{1}_{n \times 1} \bar{\mathbf{x}}$$

Now we have our zero-mean points:

$$B = X - \bar{X}$$

We want to find the eigenvector associated with the highest eigenvalue of the covariance matrix, then we want:

$$v_1 = \underset{\|v\|_2=1}{\operatorname{argmax}} (v^T B^T B v)$$

Now, we exploit the fact that by translating the points, the distance between them remain unchanged. With this in mind, we can focus our attention to the matrix B , and consider the lines that have the component $c = 0$, since we want them to pass through the centroid (which now is the origin). Our new condition becomes:

$$v^* = \operatorname{argmin}_{\|v\|_2=1} \left(\sum_i \operatorname{dist}(B_i, \operatorname{proj}_v B_i)^2 \right)$$

$$v^* = \operatorname{argmin}_{\|v\|_2=1} \left(\sum_i B_i B_i^T - (B_i \cdot v)^2 \right)$$

$$v^* = \operatorname{argmax}_{\|v\|_2=1} \left(\sum_i (B_i \cdot v)^2 \right), \text{ since } v \text{ only appears in the second term}$$

$$v^* = \operatorname{argmax}_{\|v\|_2=1} \left(\sum_i (B_i v)^2 \right) \text{ Recall that } B_i \text{ is a line vector, and } v \text{ is a column vector}$$

$$v^* = \operatorname{argmax}_{\|v\|_2=1} \left(\sum_i v^T B_i^T B_i v \right)$$

$$v^* = \operatorname{argmax}_{\|v\|_2=1} \left(v^T \left(\sum_i B_i^T B_i \right) v \right)$$

$$v^* = \operatorname{argmax}_{\|v\|_2=1} \left(v^T B^T B v \right)$$

Which is the definition of v_1 .