Problem Set 8 Solutions

Computer Vision
University of Bern
Fall 2021

1 Fitting

1. Show the procedure to fit a line $y = c_1 x + c_0$ to some observed points (x_j, y_j), $j = 1, ..., n$ by minimizing the least squares error:

$$\epsilon(c_1, c_0) = \sum_{j=1}^{n} (c_1 x_j + c_0 - y_j)^2 \tag{1}$$

Solution By differentiation we get

$$\nabla\epsilon = 2 \begin{pmatrix} \sum_{j=1}^{n} (c_1 x_j + c_0 - y_j) x_j \\ \sum_{j=1}^{n} (c_1 x_j + c_0 - y_j) \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{2}$$

$$= 2 \begin{pmatrix} c_1 \sum_{j=1}^{n} x_j^2 + c_0 \sum_{j=1}^{n} x_j - \sum_{j=1}^{n} x_j y_j \\ c_1 \sum_{j=1}^{n} x_j + c_0 \sum_{j=1}^{n} 1 - \sum_{j=1}^{n} y_j \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \tag{3}$$

At the minimum we must have $\nabla\epsilon = 0$, and this gives the equations

$$\begin{bmatrix} \sum_{j=1}^{n} x_j^2 & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j & n \end{bmatrix} \begin{bmatrix} c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} x_j y_j \\ \sum_{j=1}^{n} y_j \end{bmatrix} \tag{4}$$

2. Compute the parameters c_1 and c_0 of the best line through the points (0,-7), (2,-1) and (4,5).

Solution By substituting the given values into the normal equation we get

$$\begin{bmatrix} (0^2+2^2+4^2) & (0+2+4) \\ (0+2+4) & 3 \end{bmatrix} \begin{bmatrix} c_1 \\ c_0 \end{bmatrix} = \begin{bmatrix} (0\cdot(-7)+2\cdot(-1)+4\cdot 5) \\ (-7-1+5) \end{bmatrix} \tag{5}$$

From which we get the system

$$\begin{cases} 20c_1 + 6c_0 = 18 \\ 6c_1 + 3c_0 = -3 \end{cases} \tag{6}$$

that readily gives $c_1 = 3$ and $c_0 = -7$ and therefore $y = 3x - 7$.

3. Show that the Prewitt gradient operator can be obtained by fitting the least-squares plane through the 3×3 neighborhood of the intensity function.

   Hint: Fit a plane to the nine points $(x + \delta x, y + \delta y, I[x + \delta x, y + \delta y])$ where δx and δy range through $-1, 0, +1$. Then, having the planar model $z = ax + by + c$ that best fits the intensity surface (i.e. minimizes the least-squares error in the z-direction), show that the two Prewitt masks actually compute a and b.

Solution We denote the nine points by (x_j, y_j, z_j), $j = 1, \ldots, 9$. The least-squares error for the planar model is

$$\epsilon(a, b, c) = \sum_{j=1}^{n} (ax_j + by_j + c - z_j)^2. \tag{7}$$

By differentiation we get

$$\nabla\epsilon = 2 \begin{pmatrix} \sum_{j=1}^{n}(ax_j + by_j + c - z_j)x_j \\ \sum_{j=1}^{n}(ax_j + by_j + c - z_j)y_j \\ \sum_{j=1}^{n}(ax_j + by_j + c - z_j) \end{pmatrix} \tag{8}$$

At the minimum we must have $\nabla\epsilon = 0$, and this gives the equations

$$\begin{bmatrix} \sum_{j=1}^{n} x_j^2 & \sum_{j=1}^{n} x_j y_j & \sum_{j=1}^{n} x_j \\ \sum_{j=1}^{n} x_j y_j & \sum_{j=1}^{n} y_j^2 & \sum_{j=1}^{n} y_j \\ \sum_{j=1}^{n} x_j & \sum_{j=1}^{n} y_j & n \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} x_j z_j \\ \sum_{j=1}^{n} y_j z_j \\ \sum_{j=1}^{n} z_j \end{bmatrix} \tag{9}$$

For our 3×3 neighborhood we may fix the coordinate origin at the center of the mask. In this case x_j and y_j range through $-1, 0, +1$ and the normal equations have the form

$$\begin{bmatrix} 6 & 0 & 0 \\ 0 & 6 & 0 \\ 0 & 0 & 9 \end{bmatrix} \begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \sum_{j=1}^{n} x_j z_j \\ \sum_{j=1}^{n} y_j z_j \\ \sum_{j=1}^{n} z_j \end{bmatrix}. \tag{10}$$

Therfore the optimal plane is obtained by

$$\begin{bmatrix} a \\ b \\ c \end{bmatrix} = \begin{bmatrix} \frac{1}{6}\sum_{j=1}^{n} x_j z_j \\ \frac{1}{6}\sum_{j=1}^{n} y_j z_j \\ \frac{1}{9}\sum_{j=1}^{n} z_j \end{bmatrix}, \tag{11}$$

where the first two components are obtained by using the Prewitt masks:

$$M_x = \begin{bmatrix} -1 & 0 & 1 \\ -1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}, \quad M_y = \begin{bmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ -1 & -1 & -1 \end{bmatrix}. \tag{12}$$

4. **Bonus:** Derive the expression for the shortest distance between a line l expressed with equation $ax+by=d$ and the point (x_0, y_0) . Hint: It's the distance of a perpendicular line segment between the line l and the point (x_0, y_0).

Solution Suppose (x, y) is a point lying on line l, then the distance between (x, y) and (x_0, y_0) becomes $l(x, y) = (x - x_0)^2 + (y - y_0)^2$. Then, we can express the our task as:

$$min\ l(x, y)\ s.t.\ ax + by = d \tag{13}$$

which is a constrained optimization task. Since it's convex, we may use Lagrange multipliers method to solve it. Let's define the problem again with Lagrange multipliers notation.

$$L(x, y, \lambda) = (x - x_0)^2 + (y - y_0)^2 + \lambda(ax + by - d) \tag{14}$$

Note that, $ax+by-d$ term forces this expression to be 0 which is the same things as $ax + by = d$. After that point, we need to take the derivative of this expression with respect to x, y and λ.

$$\frac{\partial L}{\partial x} = 2(x - x_0) + \lambda a \tag{15}$$

$$\frac{\partial L}{\partial y} = 2(y - y_0) + \lambda b \tag{16}$$

$$\frac{\partial L}{\partial \lambda} = ax + by - d \tag{17}$$

Let's equate all of them to zero. Then, we get values for x, y and λ.

$$x = -\frac{\lambda a}{2} + x_0 \tag{18}$$

$$y = -\frac{\lambda b}{2} + y_0 \tag{19}$$

$$ax + by - d = 0 \tag{20}$$

Let's plug the x and y values found in the previous equations into the last equation to get the λ value.

$$-\frac{\lambda a^2}{2} + ax_0 - \frac{\lambda b^2}{2} + by_0 - d = 0 \tag{21}$$

$$\lambda = \frac{2(ax_0 + by_0 - d)}{a^2 + b^2} \tag{22}$$

Now, we may compute x and y values since we've already computed λ. Then, plug these x and y values into $l(x, y)$ function.

$$x = \frac{a(ax_0 + by_0 - d)}{a^2 + b^2} + x_0 \tag{23}$$

$$y = \frac{b(ax_0 + by_0 - d)}{a^2 + b^2} + y_0 \tag{24}$$

$$l(x, y) = \frac{a^2(ax_0 + by_0 - d)^2}{(a^2 + b^2)^2} + \frac{b^2(ax_0 + by_0 - d)^2}{(a^2 + b^2)^2} \tag{25}$$

$$l(x, y) = \frac{(ax_0 + by_0 - d)^2}{a^2 + b^2} \tag{26}$$

5. Show the procedure to fit a line $ax + by = d$ to some observed points (x_j, y_j), $j = 1, ..., n$ by minimizing the total least squares error:

$$\epsilon(a, b, d) = \sum_{j=1}^{n} (ax_j + by_j - d)^2, \tag{27}$$

where $N = (a, b)$ is the unit normal, so $\|N\|^2 = a^2 + b^2 = 1$. The line equation is $ax + by = d$. Compute the parameters a, b and d of the best line through the points (0,-7), (2,-1) and (4,5).

Solution With a and b fixed, the solution for d can be expressed as

$$d = \frac{a}{n}\sum_{i=1}^{n} x_i + \frac{b}{n}\sum_{i=1}^{n} y_i = a\bar{x} + b\bar{y}. \tag{28}$$

Substituting this to the error function, we get

$$\epsilon(a, b, d) = \sum_{j=1}^{n} (a(x_j - \bar{x}) + b(y_j - \bar{y}))^2 = \tag{29}$$

$$= \| \begin{bmatrix} x_1 - \bar{x} & y_1 - \bar{y} \\ \vdots & \vdots \\ x_n - \bar{x} & y_n - \bar{y} \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} \|^2 = N^T U^T U N. \tag{30}$$

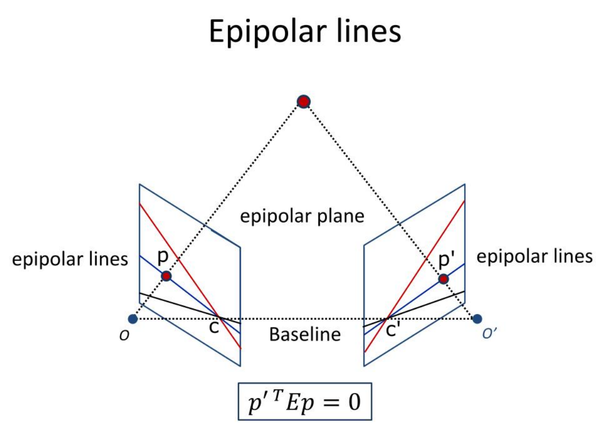
Because U^TU is symmetrical and positive semidefinite, its eigendecomposition $U^TU = V\Lambda V^T$ has nice properties. The eigenvalues λ_i (the diagonal elements of Λ) are all greater than or equal to 0 and V is orthogonal. Lets define $m = V^TN$. Because N has unit norm and V is orthogonal, m has unit norm too. The error function can be written as $\epsilon = m^T\Lambda m = m_1^2\lambda_1 + m_2^2\ \lambda_2$. From here we can see that the optimal value is $m = (0, 1)$, assuming λ_2 is the smallest eigenvalue. Therefore $N = Vm = v_2$, the eigenvector corresponding to the smallest eigenvalue.

For the 3 samples, we have $\bar{x} = 2$ and $\bar{y} = -1$.

$$U = \begin{bmatrix} -2 & -6 \\ 0 & 0 \\ 2 & 6 \end{bmatrix}, U^TU = \begin{bmatrix} 8 & 24 \\ 24 & 72 \end{bmatrix} \tag{31}$$

We can see that U has rank 1 so U^TU also has rank 1, therefore U^TU is singular. This means its smallest eigenvalue is $\lambda = 0$. This leads to $8v_x + 24v_y = 0$, which has the solution $v_x = -3\sqrt{0.1}$ and $v_y = \sqrt{0.1}$. The optimal line parameters are $a = -3\sqrt{0.1}$ and $b = \sqrt{0.1}$ and $d = a\bar{x} + b\bar{y} = -7\sqrt{0.1}$.

2 Multiview Stereo



1. The camera projection matrices of two cameras (given in the coordinate system attached to the first camera) are

$$\mathbf{C} = [\mathbf{I}\ \mathbf{0}] \text{ and } \mathbf{C}' = [\mathbf{R}\ \mathbf{t}],$$

where $\mathbf{R}$ is a rotation matrix and $\mathbf{t} = (t_1, t_2, t_3)^T$ describes the translation between the cameras. Hence, the cameras have identical internal parameters and the image points are given in the normalized image coordinates (the origin of the image coordinate frame is at the principal point and the focal length is 1).

The epipolar constraint implies that if p and p' are corresponding image points then the vectors $\overrightarrow{Op}$, $\overrightarrow{O'p'}$ and $\overrightarrow{OO'}$ are coplanar,i.e.

$$\overrightarrow{O'p'} \cdot (\overrightarrow{O'O} \times \overrightarrow{Op}) = 0, \tag{32}$$

where O and O' are the camera centers.

Let $\mathbf{p} = (x, y, 1)^T$ and $\mathbf{p'} = (x', y', 1)^T$ denote the homogeneous image coordinate vectors of p and p'. Show that the equation (32) can be written in the form

$$\mathbf{p}'^T E \mathbf{p} = 0, \tag{33}$$

where matrix $\mathbf{E}$ is the essential matrix $\mathbf{E} = [\mathbf{t}]_\times \mathbf{R}$, and

$$[\mathbf{t}]_\times = \begin{bmatrix} 0 & -t_3 & t_2 \\ t_3 & 0 & -t_1 \\ -t_2 & t_1 & 0 \end{bmatrix} \tag{34}$$

Solution In the coordinate system of the second camera we have

$$\overrightarrow{O'p'} = \mathbf{p}', \quad \overrightarrow{O'O} = \mathbf{t}, \quad \overrightarrow{Op} = \mathbf{Rp} \tag{35}$$

and by substituting these into (32) we get

$$\mathbf{p}'(\mathbf{t} \times \mathbf{Rp}) = 0, \quad \mathbf{p}'([\mathbf{t}]_\times \mathbf{Rp}) = 0 \tag{36}$$