CSE 252B: Computer Vision II

Lecturer: Serge Belongie
Scribe: Dave Berlin, Jefferson Ng

LECTURE 2
Homogeneous Linear Least Squares Problems, Two View Geometry

2.1. Introduction

We will frequently encounter problems of the form

$$A\boldsymbol{x} = \boldsymbol{0} \tag{2.1}$$

known as the Homogeneous Linear Least Squares problem. It is similar in appearance to the inhomogeneous linear least squares problem

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.2}$$

in which case we solve for $\boldsymbol{x}$ using the pseudoinverse or inverse of A. This won't work with Equation 2.1. Instead we solve it using Singular Value Decomposition (SVD), as described in the following example.

[1]Department of Computer Science and Engineering, University of California, San Diego.

March 31, 2004

2.2. Motivating Problem: Conic Fitting

Conics were mentioned in the previous lecture. Using homogeneous coordinates, they have the form

$$ax^2 + bxy + cy^2 + dxz + eyz + fz^2 = 0 \tag{2.3}$$

We can also write this as a quadratic form as

$$\boldsymbol{x}^\top C\boldsymbol{x} = 0$$

$$\boldsymbol{x} = (x, y, z)^\top$$

where C is the symmetric matrix

$$C = \begin{bmatrix} a & b/2 & d/2 \\ b/2 & c & e/2 \\ d/2 & e/2 & f \end{bmatrix}$$

As an example, for a unit circle, the corresponding matrix would be

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}$$

which corresponds to $x^2 + y^2 = z^2$. We can easily convert to inhomogeneous coordinates by letting $x' = x/z$ and $y' = y/z$, which gives the known formula for a unit circle, $(x')^2 + (y')^2 = 1$.

There are six variables in the general conic matrix C, but because of the homogenous property, they are arbitrary up to a scale factor. Therefore, there are five degrees of freedom, and five points are necessary to define a conic.

Circles are certainly conics, but they only need three points to define them. This is not a contradiction – the other two points exist, but have complex coordinates. These points are called *circular points*, that is, points where the circle crosses the line at infinity. To find their coordinates, we intersect the line at infinity $\boldsymbol{l}_\infty = (0, 0, 1)^\top$ with a circle:

$$\boldsymbol{x}^\top \boldsymbol{l}_\infty = 0$$

$$(x, y, z) \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = 0$$

and therefore z must be 0. Placing this back in the equation for a circle gives

$$x^2 + y^2 = 0$$

and therefore for the circular points, x=1, and y=$\pm$ i.

Definition 2.4. Circular points are defined as the following:

$$\boldsymbol{I} = (1, i, 0)^\top \tag{2.5}$$

$$\boldsymbol{J} = (1, -i, 0)^\top \tag{2.6}$$

All circles must pass through these points. In addition, these points are ideal points (or points at infinity) and complex conjugates of each other. Later in the course we will use these facts to help calibrate a camera.

We now formalize the conic fitting problem.

Definition 2.7. The conic fitting problem is as follows: Given $n \geq 5$ points in a plane, $\{\boldsymbol{x}^i\}_{i=1}^n$ find the coefficient vector $\boldsymbol{c} = (a, b, c, d, e, f)^\top$.

At first glance, this equation seems to have little to do with homogeneous least squares. Here, the equation is of the form $\boldsymbol{x}^\top C \boldsymbol{x} = 0$. Note that C is unknown, $\boldsymbol{x}$ is known, and 0 is a scalar. In contrast, in the equation $A\boldsymbol{x} = \boldsymbol{0}$, A is known, $\boldsymbol{x}$ is unknown, and $\boldsymbol{0}$ is a vector.

This discrepancy is addressed by collecting the unknowns into the coefficient vector $\boldsymbol{c}$. The constraint that the ith point places on $\boldsymbol{c}$ is

$$\begin{bmatrix} (x^i)^2 & x^i y^i & (y^i)^2 & x^i z^i & y^i z^i & (z^i)^2 \end{bmatrix} \boldsymbol{c} = \boldsymbol{0}$$

This six element row vector (multiplying $\boldsymbol{c}$ from the left) is known as a *carrier vector*. Now we can get to the least squares problem by "stacking" the carrier vectors into a matrix. This matrix is called the *design matrix*.

Definition 2.8. The design matrix is the matrix $A \in \mathbb{R}^{n \times 6}$ in the following equation:

$$A\boldsymbol{c} = \boldsymbol{0} \tag{2.9}$$

Now it's in the recognizable homogeneous least squares format, and we can solve it. First, compute the SVD (see MaSKS Sec. A.7) of A:

$$A = U \Sigma V^\top = \sum_{i=1}^{6} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top \tag{2.10}$$

When performed in Matlab, the singular values σ_i will be sorted in descending order, so σ_6 will be the smallest. There are three cases for the value of σ_6:

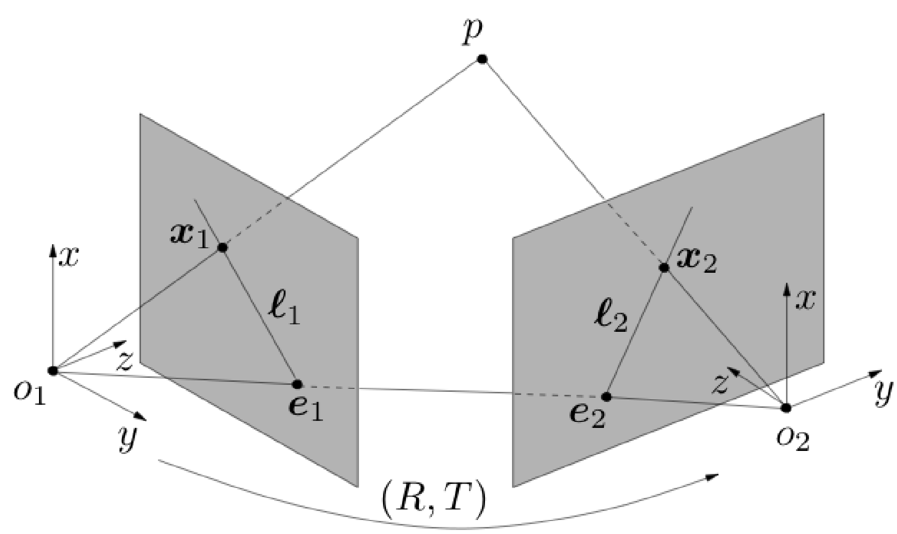
- If the conic is *exactly determined* ($n = 5$), then $\sigma_6 = 0$, and there exists a conic that fits the points exactly.
- If the conic is *overdetermined* ($n > 5$), then $\sigma_6 \geq 0$. Here σ_6 represents a "residual" or goodness of fit.
- We will not handle the case of the conic being *underdetermined* ($n < 5$).

From the SVD we take the "right singular vector" (a column from V) which corresponds to the smallest singular value, σ_6. This is the solution, $\boldsymbol{c}$, which contains the coefficients of the conic that best fits the points. We reshape this into the matrix C, and form the equation $\boldsymbol{x}^\top C\boldsymbol{x} = 0$.

To recap, note that although the expression for a conic looks nonlinear, it is only the known variables (the coordinates of the $\boldsymbol{x}^i$'s) that appear nonlinearly; we were able to write the problem in homogeneous least squares form since the coefficients appear linearly.

2.3. Two View Geometry

We now consider the geometry of two calibrated cameras viewing a scene. We assume that the cameras are related by a rigid body motion $(R, \boldsymbol{T})$. (Figure from MaSKS Ch. 5.)



Since the cameras are calibrated, we have $K_1 = K_2 = I$. The cameras are centered at o_1 and o_2, respectively. The homogeneous vectors $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the *epipoles*, and can be intuitively thought of as any of the following:

- The points where the baseline pierces the image planes
- The projection of the other camera's optical center onto each image plane
- The translation vector $\boldsymbol{T}$ (up to a scale factor)
- The direction of travel (focus of expansion)

The lines $\boldsymbol{l}_1$ and $\boldsymbol{l}_2$ are the *epipolar lines*. The plane spanned by o_1, o_2 and p is called the *epipolar plane*, and the epipolar lines are the intersections of the epipolar plane with the image planes.

2.3.1. Special Case: Rectified Stereo

Rectified stereo is the simplest case of two-view geometry in which we have two cameras that are aimed straight forward and translated horizontally w.r.t. each other, as if your eyes were looking straight ahead at something infinitely far away. In this case, the epipolar lines are horizontal, and points in one image plane map to the horizontal scan line with the same y coordinate on the other image plane.

If the cameras are not rectified in this way, how can we find corresponding points in the second image? It turns out we will still have a one-dimensional search, it just won't be as simple as being on corresponding horizontal scan lines.

2.3.2. General Two View Geometry

We specify the pose of the two cameras, g_1 and g_2 as follows:

$$g_1 = (I, \mathbf{0})$$
$$g_2 = (R, \boldsymbol{T}) \in SE(3)$$

Without loss of generality for g_1, we let its rotation and translation be the identity matrix and zero vector, respectively. For g_2, R is any rotation matrix and $\boldsymbol{T}$ is the translation vector.

A 3D point p will have coordinates $\boldsymbol{X}_1$ and $\boldsymbol{X}_2$ when viewed from g_1 and g_2 respectively. The following equation relates coordinate systems from camera 1 and camera 2.

$$\boxed{\boldsymbol{X}_2 = R\boldsymbol{X}_1 + \boldsymbol{T}} \tag{2.11}$$

2.3.3. The Epipolar Constraint and the Essential Matrix

We now want to find a relation between a point on one image and its possible locations on the other image. We begin by converting the image points into homogeneous coordinates.

For some depths λ_i we have

$$\boldsymbol{X}_1 = \lambda_1 \boldsymbol{x}_1, \quad \boldsymbol{X}_2 = \lambda_2 \boldsymbol{x}_2$$

which means

$$\lambda_2 \boldsymbol{x}_2 = R\lambda_1 \boldsymbol{x}_1 + \boldsymbol{T}$$

but λ_1, λ_2 are unknown. To solve this problem, Longuet-Higgins eliminated these depths algebraically as follows.

Take the cross product of both sides with $\boldsymbol{T}$,

$$\lambda_2 \widehat{T} \boldsymbol{x}_2 = \widehat{T} R \lambda_1 \boldsymbol{x}_1 + \underbrace{\widehat{T} T}_{=\mathbf{0}}$$

and take the inner product with $\boldsymbol{x}_2$,

$$\lambda_2 \underbrace{\boldsymbol{x}_2^\top \widehat{T} \boldsymbol{x}_2}_{=\mathbf{0}} = \boldsymbol{x}_2^\top \widehat{T} R \lambda_1 \boldsymbol{x}_1$$

$$\boldsymbol{x}_2^\top \widehat{T} R \boldsymbol{x}_1 = 0$$

$$\boxed{\boldsymbol{x}_2^\top E \boldsymbol{x}_1 = 0} \tag{2.12}$$

Equation 2.12 is a bilinear form and is called the *essential constraint* or *epipolar constraint.* It gives us a line in the image plane of camera 2 for a point in the image plane of camera 1, and vice versa.

The *essential matrix* $E = \widehat{T}R \in \mathbb{R}^{3\times3}$ compactly encodes the relative camera pose $g = (R, \boldsymbol{T})$.

Thus to map a point in one image to a line in the other using the essential matrix, we apply the following equations:

$$\boldsymbol{l}_2 \sim E\boldsymbol{x}_1$$

$$\boldsymbol{x}_2^\top \boldsymbol{l}_2 = 0 \tag{2.13}$$

Alternatively, you can go the other way:

$$\boldsymbol{l}_1 \sim E^\top \boldsymbol{x}_2$$

$$\boldsymbol{x}_1^\top \boldsymbol{l}_1 = 0 \tag{2.14}$$

where $\boldsymbol{l}_1$, $\boldsymbol{l}_2$ are epipolar lines (specified in homogeneous coordinates).

2.3.4. Extracting the Epipoles From the Essential Matrix

Note that all epipolar lines in an image plane intersect at the epipole. Equivalently, the epipole has a distance of zero from every epipolar line: $\boldsymbol{e}_2^\top \boldsymbol{l}_2 = 0, \forall \boldsymbol{x}_1$, and similarly $\boldsymbol{e}_1^\top \boldsymbol{l}_1 = 0, \forall \boldsymbol{x}_2$.

For this to hold true, $\boldsymbol{e}_2^\top E$ and $E\boldsymbol{e}_1$ must be zero vectors, i.e.,

$$\boldsymbol{e}_2^\top E = \mathbf{0}, \qquad E\boldsymbol{e}_1 = \mathbf{0}$$

Thus $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are vectors in the right and left null space of E, respectively, i.e., the left and right singular vectors of E with singular value 0.