HW2: Solving Homogeneous Systems

Owen Jow

May 5, 2019

1 Motivation

A system of m linear equations in n variables can be written in the form $\mathbf{Ax} = \mathbf{b}$, where $\mathbf{A}$ is an $m \times n$ matrix of coefficients, $\mathbf{x}$ is an $n \times 1$ vector of variables, and $\mathbf{b}$ is an $m \times 1$ vector of constant terms. If $\mathbf{b} = \mathbf{0}$, i.e. all the constant terms are zero, we call the system of equations *homogeneous*.

In HW2, homogeneous systems of equations arise when computing the fundamental matrix via the eight-point algorithm ($\mathbf{Af} = \mathbf{0}$) and when computing epipoles ($\mathbf{F}^T\mathbf{e} = \mathbf{0}$, $\mathbf{Fe}' = \mathbf{0}$). Of course, a homogeneous system always admits a trivial solution of $\mathbf{0}$, but this solution is typically useless. Obviously $\mathbf{0}$ is not a valid fundamental matrix!

In short, we would like to find the best solution for $\mathbf{Ax} = \mathbf{0}$ that isn't $\mathbf{x} = \mathbf{0}$. We will assume that $\mathbf{A}$ is rank-deficient, meaning a nontrivial solution does exist. Depending on the shape of $\mathbf{A}$, there are two main ways we can go about this...

2 Case #1: A is square

A $d \times d$ matrix maps d-dimensional vectors to d-dimensional vectors. Accordingly, it has eigenvectors [vectors which are (only) scaled under $\mathbf{A}$]. Again, we're assuming that $\mathbf{A}$ has a nontrivial null space and hence it has at least one zero eigenvalue. Then, as our solution, we can simply use an eigenvector with a zero eigenvalue. Eigenvectors cannot be zero vectors, so the triviality issue is solved.

Since the fundamental matrix $\mathbf{F} \in \mathbb{R}^{3\times3}$ is square, we can follow this approach to find the epipoles. Namely, to find $\mathbf{e}$ we can take the eigenvector of $\mathbf{F}^T$ with the minimal eigenvalue, and to find $\mathbf{e}'$ we can take the eigenvector of $\mathbf{F}$ with the minimal eigenvalue. The reason I say "minimal eigenvalue" instead of "zero eigenvalue" is that our estimate of $\mathbf{F}$ is probably only an approximation.[1] Therefore, it might not have a zero eigenvalue exactly, and the best thing we can do is adopt the eigenvector with the *smallest* eigenvalue (since it will make $\mathbf{F}^T\mathbf{e}$ or $\mathbf{Fe}'$ as close as possible to zero).

3 Case #2: A is of arbitrary shape

In general, a matrix $\mathbf{A}$ might be non-square and thus lack eigenvectors. But it *will* have an SVD:

$$\mathbf{A} = \mathbf{USV}^T$$

[1] There are multiple potential reasons for this, e.g. precision and/or noise in the correspondences used to estimate $\mathbf{F}$. We're already thinking about this during the eight-point algorithm; it's partly why we take the rank-2 approximation.

The SVD has many applications, e.g. low-rank matrix approximation. As we will see, it can also be used to find nontrivial least-squares solutions to homogeneous systems of equations $\mathbf{Ax} = \mathbf{0}$.

The least-squares objective to minimize is $\|\mathbf{Ax} - \mathbf{b}\|^2 = \|\mathbf{Ax}\|^2 = \mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$, and in order to avoid trivial solutions we will also impose the constraint that $\|\mathbf{x}\| = 1$ (the scale of $\mathbf{x}$ doesn't matter; any scaling of a solution to $\mathbf{Ax} = \mathbf{0}$ will also be a solution. We just don't want the norm to be 0).

Note that $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$ is a multivariate function (one variable for every component of $\mathbf{x}$). To minimize it subject to $\|\mathbf{x}\| = 1 \implies \mathbf{x}^T\mathbf{x} = 1$, we can introduce a Lagrange multiplier λ and set up the Lagrangian

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

Then, to find the minimum of $\mathcal{L}$, we take the gradient[2] and set it equal to $\mathbf{0}$:

$$\begin{aligned} \mathbf{A}^T\mathbf{Ax} - \lambda\mathbf{x} &= \mathbf{0} \\ \mathbf{A}^T\mathbf{Ax} &= \lambda\mathbf{x} \end{aligned}$$

It follows that $\mathbf{x}$ should be the eigenvector of $\mathbf{A}^T\mathbf{A}$ with corresponding eigenvalue λ. Since $\mathbf{A}^T\mathbf{A}$ is positive semidefinite, its eigenvalues are nonnegative. Thus $\mathbf{x}^T\mathbf{A}^T\mathbf{Ax}$ is minimized for $\mathbf{x}$ the eigenvector corresponding to the eigenvalue closest to zero. This is just the minimal eigenvalue.

As it happens, this is just the right singular vector of $\mathbf{A}$ corresponding to the smallest singular value. Why? Because in the SVD $\mathbf{A} = \mathbf{USV}^T$, the right singular vectors (columns of $\mathbf{V}$) are the same as the eigenvectors of $\mathbf{A}^T\mathbf{A}$.

Proof. If $\mathbf{A} = \mathbf{USV}^T$, then

$$\begin{aligned} \mathbf{A}^T\mathbf{A} &= (\mathbf{USV}^T)^T(\mathbf{USV}^T) \\ &= \mathbf{VSU}^T\mathbf{USV}^T \quad (\mathbf{S} \text{ is a diagonal matrix}) \\ &= \mathbf{VS}^2\mathbf{V}^T \quad (\mathbf{U} \text{ is a unitary/orthogonal matrix}) \end{aligned}$$

$\mathbf{V}$ is also a unitary/orthogonal matrix. So if we right-multiply both sides by $\mathbf{V}$, we have

$$\mathbf{A}^T\mathbf{AV} = \mathbf{VS}^2$$

$$\begin{bmatrix} | & \cdots & | \\ \mathbf{A}^T\mathbf{Av}_1 & \cdots & \mathbf{A}^T\mathbf{Av}_n \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & \cdots & | \\ \mathbf{v}_1 & \cdots & \mathbf{v}_n \\ | & \cdots & | \end{bmatrix} \begin{bmatrix} \sigma_1^2 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_n^2 \end{bmatrix}$$

$$\begin{bmatrix} | & \cdots & | \\ \mathbf{A}^T\mathbf{Av}_1 & \cdots & \mathbf{A}^T\mathbf{Av}_n \\ | & \cdots & | \end{bmatrix} = \begin{bmatrix} | & \cdots & | \\ \sigma_1^2\mathbf{v}_1 & \cdots & \sigma_n^2\mathbf{v}_n \\ | & \cdots & | \end{bmatrix}$$

Clearly, the columns of $\mathbf{V}$ $(\mathbf{v}_1, ..., \mathbf{v}_n)$ are eigenvectors of $\mathbf{A}^T\mathbf{A}$ with eigenvalues $\sigma_1^2, ..., \sigma_n^2$. Therefore, in order to obtain the eigenvector of $\mathbf{A}^T\mathbf{A}$ with the minimal eigenvalue, we can take the right singular vector $\mathbf{v}_i$ of $\mathbf{A} = \mathbf{USV}^T$ with the smallest singular value σ_i. *This is what we will do in the context of the eight-point algorithm, in which we have a potentially overdetermined system* $\mathbf{Af} = \mathbf{0}$.

[2] https://www.math.uwaterloo.ca/~hwolkowi/matrixcookbook.pdf