Lecture notes: overdetermined homogeneous linear system

Karel Zimmermann

We search for a non-trivial solution $\mathbf{x} \in \mathbb{R}^n$ of the overdetermined homogeneous linear system

$$\mathtt{A}\mathbf{x} = \mathbf{0},$$

where *non-trivial* means $\mathbf{x} \neq 0$ and *overdetermined* means that there are more independent equations than unknowns (i.e. dim rng($\mathtt{A}$) $\geq n$). Since there is no exact non-trivial solution of such overdetermined system, the solution which minimize algebraic distance $\|\mathtt{A}\mathbf{x}\|$ is searched. To avoid the trivial solution we constrain solutions on the unit sphere, i.e. $\|\mathbf{x}\| = 1$, which yields the following constrained least-squares problem

$$\mathbf{x}^* = \arg\min_{\mathbf{x}} \|\mathtt{A}\mathbf{x}\|, \text{ subject to } \|\mathbf{x}\| = 1. \tag{1}$$

This problem has closed-form solution, which is equal to the eigen-vector of $\mathtt{A}^\top\mathtt{A}$ with the smallest corresponding eigen-value (MATLAB tip: `[W D]=EIG(A'*A); x=W(:,1)`). It is the same as the singular-vector of $\mathtt{A}$ which corresponds to the smallest singular-value (MATLAB tip: `[U S V]=SVD(A); x=V(:,end)`). For the sake of completeness, derivation of this solution is provided in the next paragraph.

We solve problem (1) by introducing Lagrange function

$$\begin{aligned} L(\mathbf{x}, \lambda) &= \|\mathtt{A}\mathbf{x}\| + \lambda(1 - \|\mathbf{x}\|) = &(2)\\ &= \mathbf{x}^\top\mathtt{A}^\top\mathtt{A}\mathbf{x} + \lambda(1 - \mathbf{x}^\top\mathbf{x}). &(3)\end{aligned}$$

Critical points (i.e. points in which local extrema can be achieved) of the Lagrange function are found by equaling derivatives to zero

$$\begin{aligned} \frac{\partial L(\mathbf{x}, \lambda)}{\partial \mathbf{x}} &= 2\mathtt{A}^\top\mathtt{A}\mathbf{x} - 2\lambda\mathbf{x} = \mathbf{0} &(4)\\ \frac{\partial L(\mathbf{x}, \lambda)}{\partial \lambda} &= 1 - \mathbf{x}^\top\mathbf{x} = \mathbf{0}. &(5)\end{aligned}$$

Equation (4) is simply rewritten as the characteristic equation

$$(\mathtt{A}^\top\mathtt{A} - \lambda\mathtt{I})\mathbf{x} = \mathbf{0}, \tag{6}$$

of $\mathtt{A}^\top\mathtt{A}$. Therefore, every eigen-vector $\mathbf{x}$ of $\mathtt{A}^\top\mathtt{A}$ with corresponding eigen-values λ is critical point and the one which yields the smallest criterion value $\|\mathtt{A}\mathbf{x}\|$ of problem (1) is chosen. Using equation (6) and the constraint (5), it is shown that the criterion values in critical points are equal to corresponding eigen-values:

$$\|\mathtt{A}\mathbf{x}\| = \mathbf{x}^\top \mathtt{A}^\top \mathtt{A}\mathbf{x} = \mathbf{x}^\top \lambda \mathbf{x} = \lambda \mathbf{x}^\top \mathbf{x} = \lambda \|\mathbf{x}\| = \lambda.$$

Therefore the solution of problem (1) is the eigen-vector of $\mathtt{A}^\top \mathtt{A}$ with the smallest eigen-value.