Some Examples, Questions, and Answers: Lectures 9 – 10 – COMPSCI 369

Georgy Gimel'farb

Least Squares

Question 9-10.1: What least squares problem is to be stated and solved in order to estimate a solution of an insolvable overdetermined linear system $\mathbf{Au} = \mathbf{b}$:

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}}_{\mathbf{u}} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{b}}$$

of m equations with n unknowns $\mathbf{u}$; $m \gg n$? What equation is actually solved by the goal estimate $\mathbf{u}^*$?

Question 9-10.2: Find the non-trivial least squares solution $\mathbf{u}^* \neq \mathbf{0}$ of an insolvable overdetermined homogeneous linear system $\mathbf{Au} = \mathbf{0}$ of m homogeneous equations with n unknowns $\mathbf{u}$; $m \gg n$.

Pseudoinverse

Question 9-10.3: Find the pseudoinverse $\mathbf{A}^+$ of the 4×2 matrix $\mathbf{A} = \begin{bmatrix} 1 & -1 \\ 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}$ by using the SVD $\mathbf{A} = \mathbf{UDV}^\mathsf{T}$ and compute $\mathbf{A}^+\mathbf{A}$ and $\mathbf{AA}^+$.

Answers

To Question 9-10.1: The estimate $\mathbf{u}^*$ of the goal solution minimises the sum of squared errors $E = \sum\limits_{i=1}^{m} e_i^2 \equiv (\mathbf{b} - \mathbf{Au})^\mathsf{T} (\mathbf{b} - \mathbf{Au})$ where $e_i = b_i - \sum\limits_{j=1}^{n} a_{ij} u_j$; $i = 1, \dots, m$, are individual errors for all the m equations:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \left\{ (\mathbf{b} - \mathbf{A}\mathbf{u})^\mathsf{T} (\mathbf{b} - \mathbf{A}\mathbf{u}) \right\}$$

Analytical minimisation gives the normal equation $\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{u}^* = \mathbf{A}^\mathsf{T}\mathbf{b}$ with the square $n \times n$ matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ at the left-hand side and the n-component vector $\mathbf{A}^\mathsf{T}\mathbf{b}$ at the right-hand side. It is the normal equation that is solved tby the goal estimate $\mathbf{u}^*$.

To Question 9-10.2: In this case, the goal non-zero estimate minimises the sum of squared errors $E = (\mathbf{A}\mathbf{u})^\mathsf{T} \mathbf{A}\mathbf{u} = \mathbf{u}^\mathsf{T}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{u}$, subject to the constraint $\mathbf{u} \neq \mathbf{0}$. To exclude the trivial zero solution, let us normalise the goal estimate to keep its unit length: $\mathbf{u}^\mathsf{T}\mathbf{u} = 1$, and apply the constrained minimisation using the method of Lagrange:

$$\mathbf{u}^* = \arg\min_{\mathbf{u}} \left\{ \mathbf{u}\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{u} - \lambda(\mathbf{u}^\mathsf{T}\mathbf{u} - 1) \right\}$$

Zero first derivative of the Largrangian, $2\mathbf{A}^\mathsf{T}\mathbf{A}\mathbf{u} - 2\lambda\mathbf{u} = \mathbf{0}$, determines the vectors $\mathbf{u}^*$ as the eigenvectors of the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ and the corresponding Lagrange factors λ^* as the relevant eigenvalues. Because in this case $E = \lambda^*$, the desired solution is the eigenvector of the matrix $\mathbf{A}^\mathsf{T}\mathbf{A}$ with the smallest eigenvalue.

To Question 9-10.3: The eigenvectors and eigenvalues of the 2×2 matrix $\mathbf{A}^\mathsf{T}\mathbf{A} = \begin{bmatrix} 4 & 0 \\ 0 & 4 \end{bmatrix}$ and the 4×4 matrix $\mathbf{A}\mathbf{A}^\mathsf{T} = \begin{bmatrix} 2 & 0 & 0 & -2 \\ 0 & 2 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ -2 & 0 & 0 & 2 \end{bmatrix}$ suggest that the SVD is as follows:

$$\mathbf{A} = \underbrace{\begin{bmatrix} 0.5 & 0.5 \\ 0.5 & -0.5 \\ -0.5 & 0.5 \\ -0.5 & -0.5 \end{bmatrix}}_{\mathbf{U}} \underbrace{\begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}}_{\mathbf{D}} \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{V}^\mathsf{T}}$$

giving the following pseudoinverse (which is proportional to $\mathbf{A^T}$ in this example due to the orthogonal columns of the matrix $\mathbf{A}$):

$$\mathbf{A}^+ = \underbrace{\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{V}} \underbrace{\begin{bmatrix} 0.5 & 0 \\ 0 & 0.5 \end{bmatrix}}_{\mathbf{D}^+} \underbrace{\begin{bmatrix} 0.5 & 0.5 & -0.5 & -0.5 \\ 0.5 & -0.5 & 0.5 & -0.5 \end{bmatrix}}_{\mathbf{U}^\mathsf{T}} = \frac{1}{4}\begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$

Therefore, $\mathbf{A}^+\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\mathbf{A}\mathbf{A}^+ = \frac{1}{2}\begin{bmatrix} 1 & 0 & 0 & -1 \\ 0 & 1 & -1 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix}$