18.085 Pset #3 Solutions

August Trollbäck

July 2020

Question 1.

- If $\mathbf{A} \in \mathbb{R}^{m \times n}$ has linearly independent columns, then $\mathbf{K} = \mathbf{A}^T\mathbf{A}$ is positive definite.

  Proof: Let $\mathbf{u} \in \mathbb{R}^n$, then:

  $$\mathbf{u}^T\mathbf{K}\mathbf{u} = \mathbf{u}^T\mathbf{A}^T\mathbf{A}\mathbf{u} = (\mathbf{A}\mathbf{u})^T(\mathbf{A}\mathbf{u}) = ||\mathbf{A}\mathbf{u}||^2 \geq 0$$

  And furthermore because the columns of $\mathbf{A}$ are independent, there is no non-zero vector $\mathbf{u}$ such that $\mathbf{A}\mathbf{u} = \mathbf{0}$. This is because $\mathbf{A}\mathbf{u}$ is a linear combination of the columns of $\mathbf{A}$ (call them $\mathbf{c}_i$), and independence means there is no nontrivial solution to $u_1\mathbf{c}_1 + \cdots + u_n\mathbf{c}_n + \cdots = \mathbf{0}$. Therefore if $\mathbf{u} \neq \mathbf{0}$ then $\mathbf{A}\mathbf{u} \neq \mathbf{0}$ and so $||\mathbf{A}\mathbf{u}||^2 > 0$, which is the condition for $\mathbf{K}$ to be positive-definite.

- If $\mathbf{A} \in \mathbb{R}^{m \times n}$ has linearly independent columns and $\mathbf{C}$ is positive definite, then $\mathbf{K} = \mathbf{A}^T\mathbf{C}\mathbf{A}$ is also positive-definite.

  Proof: Similar to before we can rewrite $\mathbf{u}^T\mathbf{K}\mathbf{u}$ as $(\mathbf{A}\mathbf{u})^T\mathbf{C}(\mathbf{A}\mathbf{u})$. Define a new vector $\mathbf{v} = \mathbf{A}\mathbf{u}$ so that $\mathbf{u}^T\mathbf{K}\mathbf{u} = \mathbf{v}^T\mathbf{C}\mathbf{v}$. From before we know if $\mathbf{u} \neq \mathbf{0}$ then $\mathbf{v} \neq \mathbf{0}$, and by the definition of $\mathbf{C}$ being positive definite, for non-zero $\mathbf{v}$ we have $\mathbf{v}^T\mathbf{C}\mathbf{v} > 0$. Therefore $\mathbf{u}^T\mathbf{K}\mathbf{u} > 0$ so $\mathbf{K}$ is also positive definite.

Question 2.

Because $m < n$ for this matrix it is easier to start with $\mathbf{A}\mathbf{A}^T$

$$\mathbf{A}\mathbf{A}^T = \begin{pmatrix} 3 & 1 & 2 \\ 1 & 1 & 0 \\ 2 & 0 & 2 \end{pmatrix}$$

Find the eigenvalues of this:

$$\det(\mathbf{A}\mathbf{A}^T - \lambda\mathbf{I}) = \det\begin{pmatrix} 3-\lambda & 1 & 2 \\ 1 & 1-\lambda & 0 \\ 2 & 0 & 2-\lambda \end{pmatrix}$$
$$= (3-\lambda)(1-\lambda)(2-\lambda) - (2-\lambda) - 4(1-\lambda)$$
$$= -\lambda^3 + 6\lambda^2 - 6\lambda = -\lambda(\lambda^2 - 6\lambda + 6)$$

The roots are $\lambda_1 = 3+\sqrt{3}, \lambda_2 = 3-\sqrt{3}, \lambda_3 = 0$. We wish to find the corresponding eigenvectors. For $\lambda_3 = 0$ this is the usual nullspace, and by inspection one such vector is $(1,-1,-1)^T$. We'd like the eigenvectors to be of unit length, so let $\mathbf{u}_3 = (1/\sqrt{3}, -1/\sqrt{3}, -1/\sqrt{3})^T$. For $\mathbf{u}_1$ we solve the following, by using the bottom two rows to relate y and z to x:

$$(\mathbf{A}\mathbf{A}^T - \lambda_1\mathbf{I})\mathbf{u}_1 = \begin{pmatrix} -\sqrt{3} & 1 & 2 \\ 1 & -2-\sqrt{3} & 0 \\ 2 & 0 & -1-\sqrt{3} \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \mathbf{0}$$

$$y = (2+\sqrt{3})^{-1}x \qquad z = 2(1+\sqrt{3})^{-1}x$$

Concretely, we can let $x = 1$ and then normalize each component by $\sqrt{x^2+y^2+z^2} = 3-\sqrt{3}$ to get unit vector $\mathbf{u}_1 = (1, 1/(2+\sqrt{3}), 2/(1+\sqrt{3}))^T/(3-\sqrt{3})$. Following the same steps for λ_2 we get the unit vector $\mathbf{u}_2 = (1/(2+\sqrt{3}), 1, -2/(1+\sqrt{3}))^T/(3-\sqrt{3})$.

Then the eigenvectors of $\mathbf{A}^T\mathbf{A}$ need to be found.

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 2 & 0 & 1 \\ 2 & 2 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 2 \end{pmatrix}$$

I will skip the computation of these, but by inspection we can see both $(1,-1,0,0)$ and $(0,0,1,0)$ are eigenvectors for $\lambda = 0$ (null space). The final numerical SVD is:

$$\mathbf{A} = \begin{pmatrix} 0.789 & 0.211 & 0.577 \\ 0.211 & 0.789 & -0.577 \\ 0.577 & -0.577 & -0.577 \end{pmatrix}\begin{pmatrix} 2.175 & 0 & 0 & 0 \\ 0 & 1.126 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0.628 & -0.325 & 0.707 & 0 \\ 0.628 & -0.325 & -0.707 & 0 \\ 0 & 0 & 0 & 1 \\ 0.460 & 0.888 & 0 & 0 \end{pmatrix}^T$$

Question 3.

See Matlab solution code.

Question 4.

Define the following matrices/vectors:

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} \qquad \mathbf{x} = \begin{pmatrix} C \\ D \end{pmatrix} \qquad \mathbf{b} = \begin{pmatrix} 1 \\ 9 \\ 9 \\ 21 \end{pmatrix}$$

The least-squares solution for $\mathbf{x}$ is given by the normal equations:

$$\mathbf{A}^T\mathbf{A}\mathbf{x} = \mathbf{A}^T\mathbf{b} \implies \mathbf{x} = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{b}$$

Compute $(\mathbf{A}^T\mathbf{A})^{-1}$:

$$\mathbf{A}^T\mathbf{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 1 & 1 \\ 1 & 3 \\ 1 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 8 \\ 8 & 26 \end{pmatrix}$$

$$(\mathbf{A}^T\mathbf{A})^{-1} = \frac{1}{40} \begin{pmatrix} 26 & -8 \\ -8 & 4 \end{pmatrix}$$

Now $\mathbf{x}$ can be found:

$$\begin{aligned} \mathbf{x} &= \frac{1}{40} \begin{pmatrix} 26 & -8 \\ -8 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 3 & 4 \end{pmatrix} \begin{pmatrix} 1 \\ 9 \\ 9 \\ 21 \end{pmatrix} \\ &= \frac{1}{40} \begin{pmatrix} 26 & -8 \\ -8 & 4 \end{pmatrix} \begin{pmatrix} 40 \\ 120 \end{pmatrix} = \begin{pmatrix} 2 \\ 4 \end{pmatrix} \end{aligned}$$

So $C = 2$ and $D = 4$ are the best parameters.

Question 5.

$$f(x,y,z) = -x^2 - y^2 - z^2 + xy + yz + xz$$

We can determine the type of critical point at $x = y = z = 0$ by inspecting the Hessian of f at that point. The Hessian is:

$$\mathbf{H} = \begin{pmatrix} \partial^2 f/\partial x^2 & \partial^2 f/\partial x \partial y & \partial^2 f/\partial x \partial z \\ \partial^2 f/\partial y \partial x & \partial^2 f/\partial y^2 & \partial^2 f/\partial y \partial z \\ \partial^2 f/\partial z \partial x & \partial^2 f/\partial z \partial y & \partial^2 f/\partial z^2 \end{pmatrix} = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix}$$

We can immediately see that $(1,1,1)^T$ is in the null space of $\mathbf{H}$, so there is a zero eigenvalue. This means the critical point cannot be characterized without further information.

Question 6.

Our three elongations are

$$\begin{aligned} e_1 &= u_1 \\ e_2 &= u_2 - u_1 \\ e_3 &= u_3 - u_2 \end{aligned}$$

Our $\mathbf{A}$ matrix is then just

$$\mathbf{A} = \begin{pmatrix} 1 & & \\ -1 & 1 & \\ & -1 & 1 \end{pmatrix}$$

The $\mathbf{C}$ matrix is defined as before to give us the forces $w_i = c_i e_i$:

$$\mathbf{C} = \begin{pmatrix} c_1 & & \\ & c_2 & \\ & & c_3 \end{pmatrix}$$

So we wish to solve $\mathbf{A}^T\mathbf{C}\mathbf{A}\mathbf{u} = \mathbf{f}$. Using $\mathbf{C} = \mathbf{I}$ and $\mathbf{f} = (1,1,1)^T$ we get displacements $\mathbf{u} = (3,5,6)^T$.