COMPULSORY EXERCISE 3

JOHAN ÅMDAL ELIASSEN

1. Exercise 8.10: Fitting a circle to points

Given points $\{t_i, y_i\}_{i=1}^m \subset \mathbb{R}^2$, $m \geq 3$, we wish to fit a circle to these points. That is, we wish to satisfy

$$(t_i - c_1)^2 + (y_i - c_2)^2 = r^2 \,, \;\; 1 \leq i \leq m \,, \tag{*}$$

hence c_1, c_2 and r are the unknowns to be solved for.

a. We have

$$(*) = t_i^2 + y_i^2 + c_1^2 + c_2^2 - 2c_1 t_i - 2c_2 y_i \,, \tag{1.1}$$

or equivalently,

$$2c_1 t_i^2 + 2c_2 y_i^2 - c_1^2 - c_2^2 + r^2 = t_i^2 + y_i^2 \,. \tag{1.2}$$

Set $x_1 = 2c_1$, $x_2 = 2c_2$, and $x_3 = -c_1^2 - c_2^2 + r^2$, and this becomes

$$t_i x_1 + y_i x_2 + x_3 = t_i^2 + y_i^2 \,, \tag{1.3}$$

or in abbreviated matrix form,

$$\underbrace{\begin{pmatrix} t_i, & y_i, & 1 \end{pmatrix}}_{=:\mathbf{A}} \begin{pmatrix} x_1, & x_2, & x_3 \end{pmatrix}^T = \underbrace{\begin{pmatrix} t_i^2 + y_i^2 \end{pmatrix}}_{=:\mathbf{b}}, \;\; 1 \leq i \leq m \,. \tag{1.4}$$

If we can solve the linear system above, it is trivial to derive the original unknowns. We have $c_1 = x_1/2$, $c_2 = x_2/2$ and $r^2 = c_1^2 + c_2^2 + x_3$.

b. LSQ: minimize E given by

$$\begin{aligned} E(\mathbf{x} = ||A\mathbf{x} - b||_2^2 &= \sum_{i=1}^m \left((t_i x_1 + y_i x_2 + x_3) - (t_i^2 + y_i^2)\right)^2 \\ &= \sum_{i=1}^m \left(t_i(x_1 - t_i) + y_i(x_2 - y_i) + x_3\right)^2 \,. \end{aligned}$$

While it's by far the most unwieldy way to solve this problem, one way to proceed is to set the gradient of E to zero.

$$\frac{\partial E}{\partial x_j} = \sum_{i=1}^m 2q_{i,j}(t_i(x_1 - t_i) + y_i(x_2 - y_i) + x_3) = 0 \,, \tag{1.5}$$

with $q_{i,j} = t_i, y_i$ or 1 for $j = 1, 2, 3$ respectively. Rewritten, we obtain the three equations

$$x_1 \sum_{i=1}^m t_i^2 + x_2 \sum_{i=1}^m t_i y_i + x_3 \sum_{i=1}^m t_i = \sum_{i=1}^m (t_i^3 + y_i^2 t_i)\,, \tag{1.6}$$

$$x_1 \sum_{i=1}^m t_i y_i + x_2 \sum_{i=1}^m y_i^2 + x_3 \sum_{i=1}^m y_i = \sum_{i=1}^m \left(t_i **2 y_i + y_i^3\right)\,, \tag{1.7}$$

$$x_1 \sum_{i=1}^m t_i + x_2 \sum_{i=1}^m y_i + m x_3 = \sum_{i=1}^m \left(t_i^2 + y_i^2\right)\,. \tag{1.8}$$

In matrix form, $\mathbf{Bx} = \mathbf{c}$.

c. Two more or less obvious conditions are sufficient to ensure that $\mathbf{A}$ has full column rank: first, there must be at least three distinct points (t_i, y_i), and secondly, these three points must not all lie on the same line $y(t) = \alpha t$, for some $\alpha \in \mathbb{R}$. If these conditions are met, $\mathbf{A}$ is guaranteed to have three linearly independed rows, i.e. a row rank of 3, which implies full column rank. These conditions are necessary even if the system is otherwise overdetermined.

d. Let $\{(t_i, y_i)\}_{i=1}^m = \{(1, 4), (3, 2), (1, 0)\}$, for $m = 3$. Three points on a circle uniquely determines it, so we may simply plug the points into $\mathbf{A}$ to obtain $\mathbf{x} = (2, 4, -1)^T$, or $c_1 = 1$, $c_2 = 2$, $r = 2$.
To verify that the calculations in (b) were correct, inserting the numbers into $\mathbf{B}$ and $\mathbf{c}$ as obtained there, I get

$$\begin{pmatrix} 11 & 10 & 5 \\ 10 & 20 & 6 \\ 5 & 6 & 3 \end{pmatrix} \mathbf{x} = \begin{pmatrix} 57 \\ 94 \\ 31 \end{pmatrix}\,, \tag{1.9}$$

which yields the same results.

Exercise 8.24

Given $\mathbf{A} \in \mathbb{C}^{m \times n}$, from previous exercises (8.17, 8.18), we have that $\mathbf{A}^\dagger$ is the unique matrix satisfying the following:

1. $(\mathbf{A}\mathbf{A}^\dagger)^* = \mathbf{A}\mathbf{A}^\dagger$,
2. $(\mathbf{A}^\dagger\mathbf{A})^* = \mathbf{A}^\dagger\mathbf{A}$,
3. $\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger = \mathbf{A}^\dagger$,
4. $\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \mathbf{A}$.

Hence, for a matrix $\mathbf{B}$, to show that a matrix $\mathbf{C}$ is the generalized inverse of $\mathbf{B}$ is to show that $\mathbf{B}$ and $\mathbf{C}$ together satisfy the four identities above.

a. Show that $(\mathbf{A}^*)^\dagger = (\mathbf{A}^\dagger)^*$: For notational simplicity, let $\mathbf{B} = \mathbf{A}^*$, $\mathbf{C} = (\mathbf{A}^\dagger)^*$.
(1) and (2) trivially hold; for (3) we have

$$\mathbf{CBC} = ((\mathbf{CBC})^*)^* = (\mathbf{C}^*\mathbf{B}^*\mathbf{C}^*)^* = (\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger)^* = (\mathbf{A}^\dagger)^* = \mathbf{C}\,,$$

and likewise for (4),

$$\mathbf{BCB} = ((\mathbf{BCB})^*)^* = (\mathbf{A}\mathbf{A}^\dagger\mathbf{A})^* = \mathbf{A}^* = \mathbf{B}\,. \tag{1.10}$$

Hence $\mathbf{C} = (\mathbf{A}^\dagger)^*$ is indeed the generalized inverse of $\mathbf{B} = \mathbf{A}^*$. □

b. Show that $(\mathbf{A}^\dagger)^\dagger = \mathbf{A}$.
(3) and (4) follow from (4) and (3) for $\mathbf{A}$. For (1) we have:

$$(\mathbf{A}^\dagger\mathbf{A})^* = \mathbf{A}^*(\mathbf{A}^\dagger)^*\,.$$

Now, from (a) this becomes

$$(...) = \mathbf{A}^*(\mathbf{A}^*)^\dagger = (\mathbf{A}^*(\mathbf{A}^*)^\dagger)^* = \mathbf{A}^\dagger\mathbf{A}\,.$$

again using (a).
The procedure for showing (2) is practically identical:

$$(\mathbf{A}\mathbf{A}^\dagger)^* = (\mathbf{A}^\dagger)^*\mathbf{A}^* \; = (\mathbf{A}^*)^\dagger\mathbf{A}^* = ((\mathbf{A}^*)^\dagger\mathbf{A}^*)^* = \mathbf{A}\;\mathbf{A}^\dagger,.$$

□

c. Show: $(\alpha\mathbf{A})^\dagger = \frac{1}{\alpha}\mathbf{A}^\dagger$, $\alpha \in \mathbb{C} \setminus \{0\}$.
This is more or less trivial. For (1), we have $(\alpha\mathbf{A}\frac{1}{\alpha}\mathbf{A}^\dagger)^* = (\mathbf{A}\mathbf{A}^\dagger)^* = \mathbf{A}\mathbf{A}^\dagger = \alpha\mathbf{A}\frac{1}{\alpha}\mathbf{A}^\dagger$, and similarly for (2), and for (3) and (4),

$$\frac{1}{\alpha}\mathbf{A}^\dagger\alpha\mathbf{A}\frac{1}{\alpha}\mathbf{A}^\dagger = \frac{1}{\alpha}(\mathbf{A}^\dagger\mathbf{A}\mathbf{A}^\dagger) = \frac{1}{\alpha}\mathbf{A}^\dagger\,, \tag{1.11}$$

and

$$\alpha\mathbf{A}\frac{1}{\alpha}\mathbf{A}^\dagger\alpha\mathbf{A} = \alpha\mathbf{A}\mathbf{A}^\dagger\mathbf{A} = \frac{1}{\alpha}\mathbf{A}\,. \tag{1.12}$$

□