E9-252: Mathematical Methods and Techniques in Signal Processing
Homework 5 Solutions

Instructor: Shayan Garani Srinivasa
Solutions Prepared By: Chaitanya Kumar Matcha

5 December 2016

Problem 14.2.3
Using block matrix representation, we can write

$$\boldsymbol{A}\underline{x} + \underline{b} = \begin{bmatrix}\boldsymbol{A} & \underline{b}\end{bmatrix}\begin{bmatrix}\underline{x} \\ 1\end{bmatrix}$$

Part a:
We want

$$\begin{bmatrix}\boldsymbol{A} & \underline{b}\end{bmatrix}\begin{bmatrix}\underline{x}_i^{[0]} \\ 1\end{bmatrix} = \underline{x}_i^{[1]}, \quad i = 0, 1, 2, \cdots, k. \tag{1}$$

Stacking the equations horizontally, we have

$$\underbrace{\begin{bmatrix}\boldsymbol{A} & \underline{b}\end{bmatrix}}_{\boldsymbol{T}}\underbrace{\begin{bmatrix}\underline{x}_0^{[0]} & \underline{x}_1^{[0]} & \cdots & \underline{x}_k^{[0]} \\ 1 & 1 & \cdots & 1\end{bmatrix}}_{\boldsymbol{X}^{[0]}} = \underbrace{\begin{bmatrix}\underline{x}_0^{[1]} & \underline{x}_1^{[1]} & \cdots & \underline{x}_k^{[1]}\end{bmatrix}}_{\boldsymbol{X}^{[1]}}$$

i.e.,

$$\boldsymbol{T}\boldsymbol{X}^{[0]} = \boldsymbol{X}^{[1]}. \tag{2}$$

Since the vertices are from $\mathbb{R}^2$, the matrices $\boldsymbol{T}$, $\boldsymbol{X}^{[0]}$ and $\boldsymbol{X}^{[1]}$ are of dimensions 2×3, $3 \times (k+1)$ and $2 \times (k+1)$ respectively.

$\boldsymbol{A}$ and $\underline{b}$ can be obtained from (2) by inverting the matrix $\boldsymbol{X}^{[0]}$:

$$\boldsymbol{T} = \begin{bmatrix}\boldsymbol{A} & \underline{b}\end{bmatrix} = \boldsymbol{X}^{[1]}\left(\boldsymbol{X}^{[0]}\right)^{-1}.$$

For this we need $k+1 = 3$ and the matrix $\boldsymbol{X}^{[0]}$ must be non-singular.

Therefore, $k+1 = 3$ vertices are necessary to uniquely define the affine transformation.

Part b:

If fewer vertices are available, then (1) is an under-determined set of equations. Therefore, there would be more than one affine transformation that achieves the desired mapping of vertices.

If more vertices are available, then (1) is an over-determined set of equations. In this case, $\boldsymbol{A}$ and $\underline{b}$ can be obtained using first 3 vertices $\underline{x}_0^{[0]}$, $\underline{x}_1^{[0]}$ and $\underline{x}_2^{[0]}$. If the remaining set of points $\underline{x}_3^{[1]}, \cdots \underline{x}_k^{[1]}$ can be obtained using the affine transformation $\boldsymbol{A}\underline{x} + \underline{b}$, then a solution exists. If the remaining set of points $\underline{x}_3^{[1]}, \cdots \underline{x}_k^{[1]}$ cannot be obtained using the affine transformation $\boldsymbol{A}\underline{x} + \underline{b}$, then no valid affine transformation exists. In this case, a psuedo-inverse of $\boldsymbol{X}^{[0]}$ can be used to obtain a least-squared solution.

Part c:

When 3 vertices are available, the unique transformation can be obtained as

$$\boldsymbol{T} = \begin{bmatrix} \boldsymbol{A} & \underline{b} \end{bmatrix} = \boldsymbol{X}^{[1]} \left(\boldsymbol{X}^{[0]}\right)^{-1}.$$

Part d:

We have

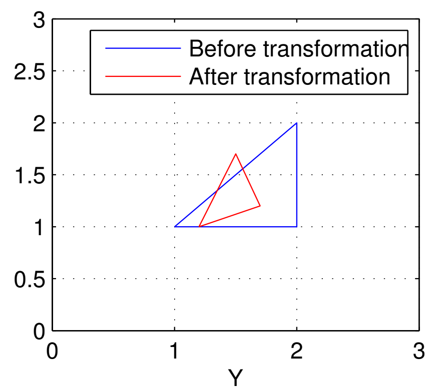
$$\boldsymbol{X}^{[0]} = \begin{bmatrix} 1 & 2 & 2 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{bmatrix}; \qquad \boldsymbol{X}^{[1]} = \begin{bmatrix} 1.2 & 1.7 & 1.5 \\ 1 & 1.2 & 1.7 \end{bmatrix};$$

Therefore, the transformation is

$$\begin{bmatrix} \boldsymbol{A} & \underline{b} \end{bmatrix} = \boldsymbol{X}^{[1]} \left(\boldsymbol{X}^{[0]}\right)^{-1} = \begin{bmatrix} 0.5 & -0.2 & 0.9 \\ 0.2 & 0.5 & 0.3 \end{bmatrix}$$

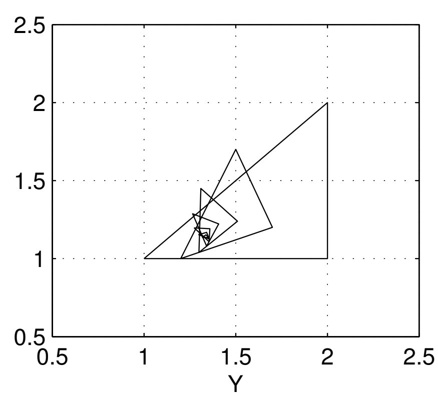
$$\implies \boldsymbol{A} = \begin{bmatrix} 0.5 & -0.2 \\ 0.2 & 0.5 \end{bmatrix}; \quad \underline{b} = \begin{bmatrix} 0.9 \\ 0.3 \end{bmatrix};$$

Following figure shows the polygon before and after the affine transformation.



Part e:

Following figure shows the ploygon after multiple iterations of the affine transformation on the given polygon.



Problem 14.2.4
We have

$$\boldsymbol{A} = \begin{bmatrix} 1 & a_{12} & \cdots & a_{1n} \\ a_{21} & 1 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 1 \end{bmatrix}.$$

Since $\boldsymbol{A}$ is diagonally dominant, we have

$$1 = |a_{i,i}| > \sum_{j \neq i} |a_{i,j}| \quad \forall i = 1, 2, \cdots, n. \tag{3}$$

$$\boldsymbol{A} - \boldsymbol{I} = \begin{bmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & 0 \end{bmatrix}.$$

Therefore, the max-norm of $\boldsymbol{A} - \boldsymbol{I}$ is

$$\|\boldsymbol{A} - \boldsymbol{I}\|_\infty = \max_i \sum_{j \neq i} |a_{i,j}| < 1 \quad \text{(From (3))}.$$

Problem 14.2-5
We have

$$\underline{x}^{[k+1]} = \boldsymbol{A}\underline{x}^{[k]} + \underline{b}. \tag{4}$$

Part a:
We will prove the following result using induction:

$$\underline{x}^{[k]} = \sum_{j=0}^{k-1} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^k \underline{x}^{[0]} \tag{5}$$

For $k = 1$, (5) is same as

$$\underline{x}^{[1]} = \underline{b} + \boldsymbol{A}\underline{x}^{[0]}$$

which is true from (5).
Assume that (5) is true for some $k = n$ i.e.,

$$\begin{aligned}
\underline{x}^{[n]} &= \sum_{j=0}^{n-1} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^n \underline{x}^{[0]} \\
\implies \underline{x}^{[n+1]} &= \boldsymbol{A}\underline{x}^{[n]} + \underline{b} \quad \text{(using (4))} \\
&= \boldsymbol{A}\left(\sum_{j=0}^{n-1} \boldsymbol{A}^j \underline{b}\right) + \boldsymbol{A}\left(\boldsymbol{A}^n \underline{x}^{[0]}\right) + \underline{b} \\
&= \sum_{j=1}^{n} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^{n+1} \underline{x}^{[0]} + \boldsymbol{A}^0 \underline{b} \\
\underline{x}^{[n+1]} &= \sum_{j=0}^{n} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^{n+1} \underline{x}^{[0]}.
\end{aligned}$$

Therefore, by mathematical induction (5) is true for all k.
From (5),

$$\begin{aligned}
(\boldsymbol{A} - \boldsymbol{I})\,\underline{x}^{[k]} &= \boldsymbol{A}\underline{x}^{[k]} - \underline{x}^{[k]} \\
&= \boldsymbol{A}\left(\sum_{j=0}^{k-1} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^k \underline{x}^{[0]}\right) - \left(\sum_{j=0}^{k-1} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^k \underline{x}^{[0]}\right) \\
&= \sum_{j=1}^{k} \boldsymbol{A}^j \underline{b} - \sum_{j=0}^{k-1} \boldsymbol{A}^j \underline{b} + \boldsymbol{A}^{k+1} \underline{x}^{[0]} - \boldsymbol{A}^k \underline{x}^{[0]} \\
&= \boldsymbol{A}^k \underline{b} - \underline{b} + \boldsymbol{A}^{k+1} \underline{x}^{[0]} - \boldsymbol{A}^k \underline{x}^{[0]} \\
(\boldsymbol{A} - \boldsymbol{I})\,\underline{x}^{[k]} &= \left(\boldsymbol{A}^k - \boldsymbol{I}\right) \underline{b} + (\boldsymbol{A} - \boldsymbol{I})\,\boldsymbol{A}^k \underline{x}^{[0]} \\
\implies \underline{x}^{[k]} &= (\boldsymbol{A} - \boldsymbol{I})^{-1} \left(\boldsymbol{A}^k - \boldsymbol{I}\right) \underline{b} + (\boldsymbol{A} - \boldsymbol{I})^{-1} (\boldsymbol{A} - \boldsymbol{I})\,\boldsymbol{A}^k \underline{x}^{[0]} \\
\therefore \underline{x}^{[k]} &= (\boldsymbol{A} - \boldsymbol{I})^{-1} \left(\boldsymbol{A}^k - \boldsymbol{I}\right) \underline{b} + \boldsymbol{A}^k \underline{x}^{[0]} \qquad (6)
\end{aligned}$$

Parts b, c:
Given $\|\boldsymbol{A}\|_2 = 1$. Therefore the largest (in magnitude) eigen value of $\boldsymbol{A}^H \boldsymbol{A}$ has magnitude 1.

$$\begin{aligned}
\boldsymbol{A} &= \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e & \lambda \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} e\cos\theta & \lambda\cos\theta - \sin\theta \\ e\sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix} \\
\implies \boldsymbol{A}^H\boldsymbol{A} &= \begin{bmatrix} e^* & 0 \\ \lambda^* & 1 \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} e & \lambda \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} e^* & 0 \\ \lambda^* & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} e & \lambda \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} e^* & 0 \\ \lambda^* & 1 \end{bmatrix} \begin{bmatrix} e & \lambda \\ 0 & 1 \end{bmatrix} \\
&= \begin{bmatrix} |e|^2 & e^*\lambda \\ e\lambda^* & 1+|\lambda|^2 \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\det\left(\boldsymbol{A}^H\boldsymbol{A} - \boldsymbol{I}x\right) &= 0 \\
\implies \det\begin{bmatrix} |e|^2 - x & e^*\lambda \\ e\lambda^* & 1+|\lambda|^2 - x \end{bmatrix} &= 0 \\
\implies \left(1+|\lambda|^2 - x\right)\left(|e|^2 - x\right) - |e|^2|\lambda|^2 &= 0 \\
\implies |e|^2 - x|e|^2 - x - |\lambda|^2 x + x^2 &= 0 \\
\implies x^2 - x\left(1+|\lambda|^2+|e|^2\right) + |e|^2 &= 0
\end{aligned}$$

$$\Delta = \left(1+|\lambda|^2+|e|^2\right)^2 - 4|e|^2 = \left(1+|\lambda|^2-|e|^2\right)^2 + 4|\lambda|^2|e|^2 > 0$$

Therefore, the roots are real.

Given $\|\boldsymbol{A}\|_2 = 1$ and since the roots are real, $x = -1$ or $x = 1$ must be a root of the characteristic equation.

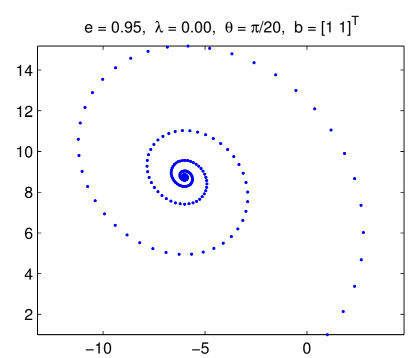
Substituting $x = -1$ gives $2 + |\lambda|^2 + 2|e|^2 = 0$ which is not possible. Therefore $x = 1$ is a root of the above equation $\implies 1 - \left(1+|\lambda|^2+|e|^2\right) - |e|^2 = 0 \implies \lambda = 0$. Using $\lambda = 0$, the characteristic equation becomes

$$x^2 - \left(1+|e|^2\right)x + |e|^2 = 0.$$

Therefore, $|e|^2$ is the other eigen value of $\boldsymbol{A}^H\boldsymbol{A}$. Since $\|\boldsymbol{A}\|_2 = 1$, $\left|e^2\right| \le 1$.

$$\lambda{=}0 \implies \boldsymbol{A} = \begin{bmatrix} e\cos\theta & \lambda\cos\theta - \sin\theta \\ e\sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix} = \begin{bmatrix} e\cos\theta & -\sin\theta \\ e\sin\theta & \cos\theta \end{bmatrix}.$$

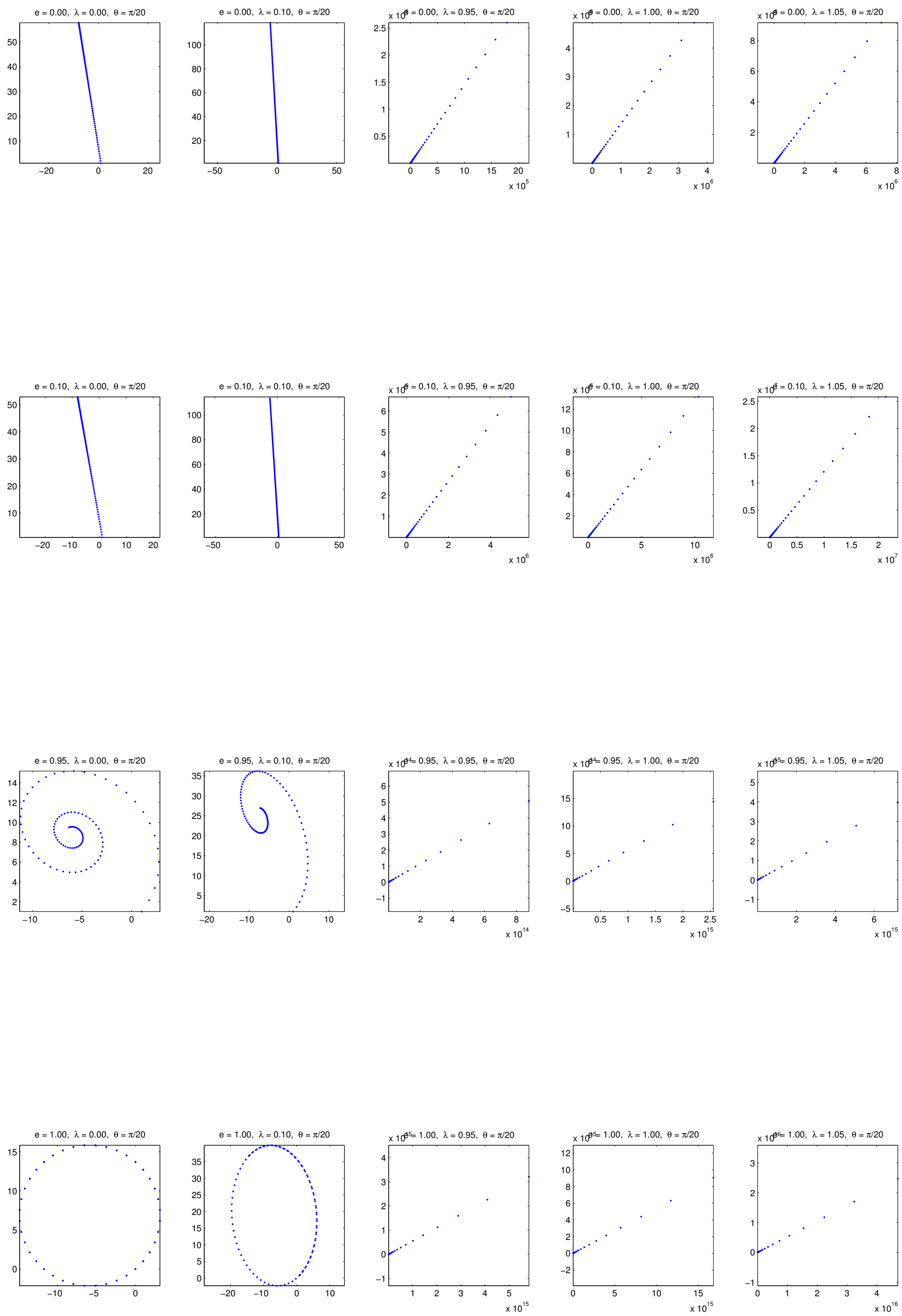
However, following plot shows that the orbit is a spiral (and not ellipse) for $\lambda = 0$.

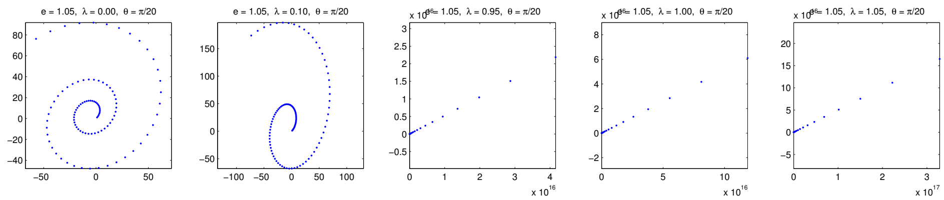


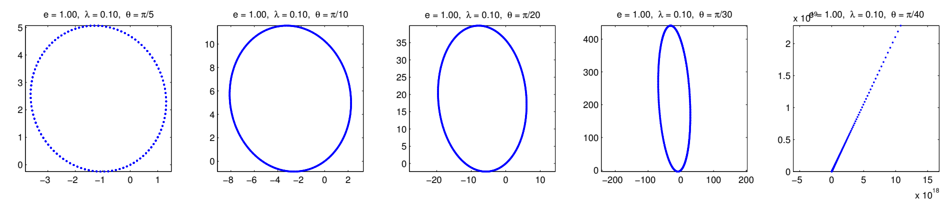
This contradicts the claim in the problem statement that $\|\boldsymbol{A}\|_2 = 1$ results in an elliptical orbit.

Part d:

Following plots show the orbits for different values of e, λ, θ for $\underline{b} = \underline{x}^{[0]} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.







Observations:

1. The orbit becomes larger with an increase in λ.
2. $e < 1$: The orbit is a spiral that converges to a point.
3. $e = 1$: The orbit is an ellipse.
4. $e > 1$: The orbit is a spiral that diverges to infinity.

Remark: The orbit is an ellipse if $e = 1$. This is because both the eigen values of $\boldsymbol{A}$ are of unit magnitude. This results in 'oscillatory' behavior of $\boldsymbol{A}^k$ as k increases.

For parts c and d of the problem, we assume $e = 1$ instead of $\|\boldsymbol{A}\|_2 = 1$ and obtain the center $\underline{x}_0$ and $\boldsymbol{U}$ for the general form of ellipse given by

$$(\underline{x} - \underline{x}_0)^T \boldsymbol{U} (\underline{x} - \underline{x}_0) = c. \tag{7}$$

For $e = 1$, we have

$$\boldsymbol{A} = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} 1 & \lambda \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} \cos\theta & \lambda\cos\theta - \sin\theta \\ \sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix}.$$

We have

$$\begin{aligned} \underline{x}^{[k]} &= \boldsymbol{A}\underline{x}^{[k-1]} + \underline{b} \\ \implies \underline{x}^{[k]} &= \boldsymbol{A}\underline{x}^{[k-1]} - \boldsymbol{A}(I - \boldsymbol{A})^{-1}\underline{b} + (I - \boldsymbol{A})^{-1}\underline{b} \\ \underline{x}^{[k]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b} &= \boldsymbol{A}\left(\underline{x}^{[k]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right) \end{aligned}$$

$$\implies \left(\underline{x}^{[k]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right)^T \boldsymbol{U}\left(\underline{x}^{[k]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right) = \left(\underline{x}^{[k-1]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right)^T \boldsymbol{A}^H\boldsymbol{U}\boldsymbol{A}\left(\underline{x}^{[k-1]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right)$$

Setting $\boldsymbol{A}^H\boldsymbol{U}\boldsymbol{A} = \boldsymbol{U}$ and comparing this to the equation of an ellipse, we have the center given by

$$\underline{x}_0 = (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}.$$

Transposing the equation (7), we can easily see that $\boldsymbol{U} = \boldsymbol{U}^T$. Also, the equation of the ellipse will still have the same form even if we scale the matrix $\boldsymbol{U}$. Therefore, we can set one of the elements of $\boldsymbol{U}$ to 1. The matrix $\boldsymbol{U} = \begin{bmatrix} 1 & u \\ u & v \end{bmatrix}$ can be obtained by solving

$$\boldsymbol{U} = \boldsymbol{A}^H\boldsymbol{U}\boldsymbol{A}$$

$$\begin{aligned}
\implies \begin{bmatrix} 1 & u \\ u & v \end{bmatrix} &= \begin{bmatrix} \cos\theta & \sin\theta \\ \lambda\cos\theta - \sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix} \begin{bmatrix} 1 & u \\ u & v \end{bmatrix} \begin{bmatrix} \cos\theta & \sin\theta \\ \lambda\cos\theta - \sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix} \\
&= \begin{bmatrix} \cos\theta & \sin\theta \\ \lambda\cos\theta - \sin\theta & \lambda\sin\theta + \cos\theta \end{bmatrix} \begin{bmatrix} \cos\theta + u\lambda\cos\theta - u\sin\theta & \sin\theta + u\lambda\sin\theta + u\cos\theta \\ u\cos\theta + v\lambda\cos\theta - v\sin\theta & u\sin\theta + v\lambda\sin\theta + v\cos\theta \end{bmatrix}
\end{aligned}$$

$$\begin{aligned}
\implies 1 &= \cos^2\theta + u\left(\lambda\cos^2\theta\right) + v\left(\lambda\cos\theta\sin\theta - \sin^2\theta\right) \\
u &= \sin\theta\cos\theta + u\left(1 + \lambda\sin\theta\cos\theta\right) + v\left(\lambda\sin^2\theta + \sin\theta\cos\theta\right) \\
v &= \lambda\sin\theta\cos\theta - \sin^2\theta + u\left(\lambda^2\sin\theta\cos\theta + \lambda\cos^2\theta\right) + v\left(\lambda^2\sin^2\theta + 2\lambda\sin\theta\cos\theta + \cos^2\theta\right)
\end{aligned}$$

Solving the above set of equations will result in

$$\begin{aligned} u &= \tan\theta \\ v &= -1. \end{aligned}$$

Therefore, $\boldsymbol{U} = \begin{bmatrix} 1 & \tan\theta \\ \tan\theta & -1 \end{bmatrix}$. Scaling by $\cos\theta$, we have

$$\boldsymbol{U} = \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}.$$

The constant c depends on the initial position $\underline{x}^{[0]}$. Therefore the equation of the ellipse is

$$\left(\underline{x} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right)^T \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \left(\underline{x} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right) = \left(\underline{x}^{[0]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right)^T \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix} \left(\underline{x}^{[0]} - (\boldsymbol{I} - \boldsymbol{A})^{-1}\underline{b}\right).$$