Mathematical Methods in Engineering and Applied Science
Problem Set 6.

Kovalev Vyacheslav

(1). Given the data:

x_i	0	1	2	3	5
y_i	1	3	3	4	6

(a) Find the best linear fit by solving the 2×2 normal system by hand $A^T A u = A^T b$, with $A = [x^T\ 1]$ and $b = y^T$. Plot the data and the fit.

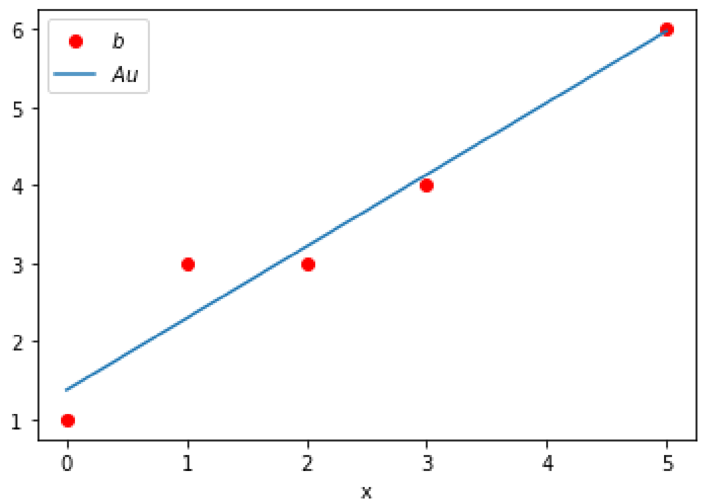
$$A = \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix};\ b = \begin{bmatrix} 1 \\ 3 \\ 3 \\ 4 \\ 6 \end{bmatrix};$$

$$A^T A u = A^T b \ \Rightarrow u = (A^T A)^{-1} A^T b = b$$

$$u = \left(\begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 \\ 3 \\ 3 \\ 4 \\ 6 \end{bmatrix} = \begin{bmatrix} 39 & 11 \\ 11 & 5 \end{bmatrix}^{-1} \begin{bmatrix} 51 \\ 17 \end{bmatrix} =$$

$$= \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix} \begin{bmatrix} 51 \\ 17 \end{bmatrix} = \frac{1}{74} \begin{bmatrix} 68 \\ 102 \end{bmatrix}$$

Data and the fit:



(b) Calculate the Moore-Penrose pseudo-inverse A^+ of A directly from its definition.

$$A^+ = (A^T A)^{-1} A^T = \left(\begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \\ 3 & 1 \\ 5 & 1 \end{bmatrix} \right)^{-1} \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} =$$

$$= \frac{1}{74} \begin{bmatrix} 5 & -11 \\ -11 & 39 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 & 3 & 5 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \frac{1}{74} \begin{bmatrix} -11 & -6 & -1 & 4 & 14 \\ 39 & 28 & 17 & 6 & -16 \end{bmatrix}$$

(c) Write down the SVD of A^+.

$A^+ = V \Sigma^+ U^T$ Let's find $A = U \Sigma V^T$

$A^TA = \begin{bmatrix} 39 & 11 \\ 11 & 5 \end{bmatrix}$ => $\lambda_{1,2} = 22 \pm \sqrt{410}$; omit process of searching e-values, e-vectors because of difficulties.

$$V = \begin{bmatrix} \frac{\sqrt{2}(17-\sqrt{410})}{2\sqrt{410-17\sqrt{410}}} & \frac{\sqrt{2}(17+\sqrt{410})}{2\sqrt{17\sqrt{410}+410}} \\ \frac{11\sqrt{2}}{2\sqrt{410-17\sqrt{410}}} & \frac{11\sqrt{2}}{2\sqrt{17\sqrt{410}+410}} \end{bmatrix}$$

$$AA^T = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 2 & 3 & 4 & 6 \\ 1 & 3 & 5 & 7 & 11 \\ 1 & 4 & 7 & 10 & 16 \\ 1 & 6 & 11 & 16 & 26 \end{bmatrix} => U = \begin{bmatrix} \frac{11}{2\sqrt{7995-392\sqrt{410}}} & \frac{11}{2\sqrt{392\sqrt{410}+7995}} \\ \frac{14-\frac{\sqrt{410}}{2}}{\sqrt{7995-392\sqrt{410}}} & \frac{\frac{\sqrt{410}}{2}+14}{\sqrt{392\sqrt{410}+7995}} \\ \frac{\frac{45}{2}-\sqrt{410}}{\sqrt{7995-392\sqrt{410}}} & \frac{\sqrt{410}+\frac{45}{2}}{\sqrt{392\sqrt{410}+7995}} \\ \frac{31-\frac{3\sqrt{410}}{2}}{\sqrt{7995-392\sqrt{410}}} & \frac{\frac{3\sqrt{410}}{2}+31}{\sqrt{392\sqrt{410}+7995}} \\ \frac{48-\frac{5\sqrt{410}}{2}}{\sqrt{7995-392\sqrt{410}}} & \frac{48+\frac{5\sqrt{410}}{2}}{\sqrt{392\sqrt{410}+7995}} \end{bmatrix}$$

$$\Sigma^+ = \begin{bmatrix} \frac{1}{\sqrt{22-\sqrt{410}}} & 0 \\ 0 & \frac{1}{\sqrt{22+\sqrt{410}}} \end{bmatrix}$$

$A^+ = V\Sigma^+U^T$ – is SVD of A^+, all matrices obtained above.

(d) What is the error vector e of the approximation and its 2-norm?

$$e = b - Au = \frac{1}{74}\begin{bmatrix} -28 \\ 52 \\ -16 \\ -10 \\ 2 \end{bmatrix}$$

$$|e|_2 = \sqrt{\frac{26}{37}}$$

(2). Find the best plane in R^3, in the least-squares sense, through the data given in the table:

x_i	1	1	2	3	5
y_i	5	3	4	10	7
z_i	2	1	2	5	5

What is the error vector and its norm?

Plane equation $ax + by + c = z$ in terms of matrices:

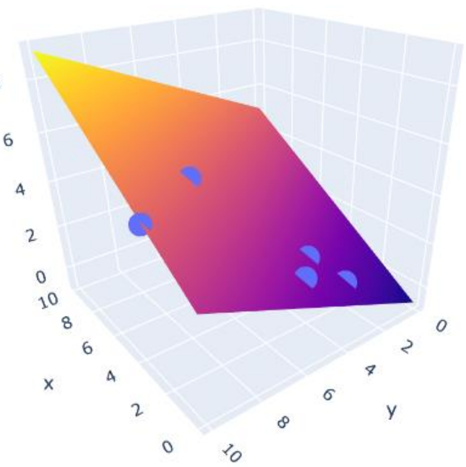
$Au = Z$; $A = [x_i \quad y_i \quad 1]$ where x_i, y_i, are columns.

$u = [a\ b\ c]$

$u = A^+Z$ where $A^+ = (A^TA)^{-1}A^T$ following the same procedure as in previous task obtain:

$$A^+ = \begin{bmatrix} -\frac{34}{215} & -\frac{56}{1075} & \frac{41}{1075} & -\frac{147}{1075} & \frac{332}{1075} \\ \frac{7}{215} & -\frac{77}{1075} & -\frac{78}{1075} & \frac{201}{1075} & -\frac{81}{1075} \\ \frac{84}{215} & \frac{796}{1075} & \frac{569}{1075} & -\frac{598}{1075} & -\frac{112}{1075} \end{bmatrix}$$

Finally: $u = \begin{bmatrix} \frac{611}{1075} \\ \frac{437}{1075} \\ -\frac{776}{1075} \end{bmatrix}$



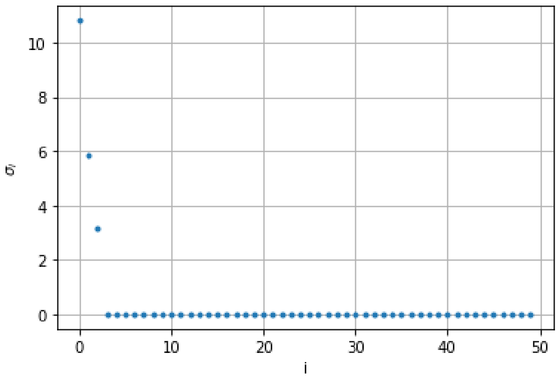
Blue points are data points.
Plane is constructed by $A'u$ for $A' = [x_i, y_i, 1]$
$[x_i] = [0, \ldots, 11], [y_i] = [0, \ldots, 11]$

$$e = Au - Z = \frac{1}{1075} \begin{bmatrix} -130 \\ 71 \\ 44 \\ 52 \\ -37 \end{bmatrix}$$

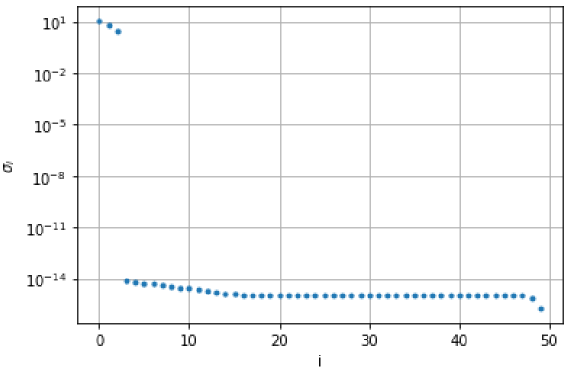
$$|e|_2 = \frac{\sqrt{1118}}{215} \approx 0.16$$

(3). Determine the dominant modes in the functionfof spacexand timet:
$f(x,t) = e^{-x^2} \sin(x + 3t) \cos(x - t)$, considering the interval $x \in [-5,5], t \in [0,10]$.

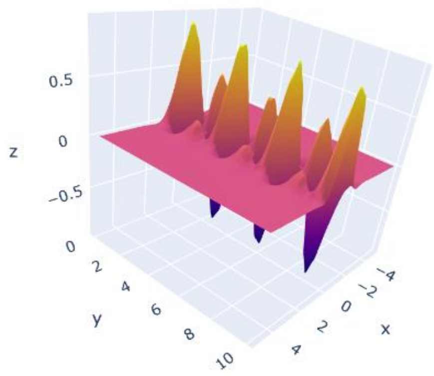
(a) Plot the singular values in uniform as well as semilog scales.
$x_i \in [-5,5], t_j \in [0,10]$. Where $i = 1 \ldots 100, j = 1 \ldots 50$
Though we got $f_{ij} = f(x_i, t_j)$
Uniform:



Semilog



(b) Plot the solution in the $x - t$ plane over the given interval.



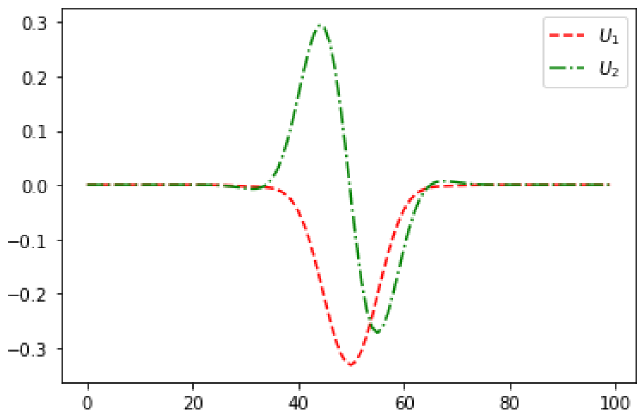
(c) How much "energy" of the solution is contained in mode 1 and in modes 1+2?

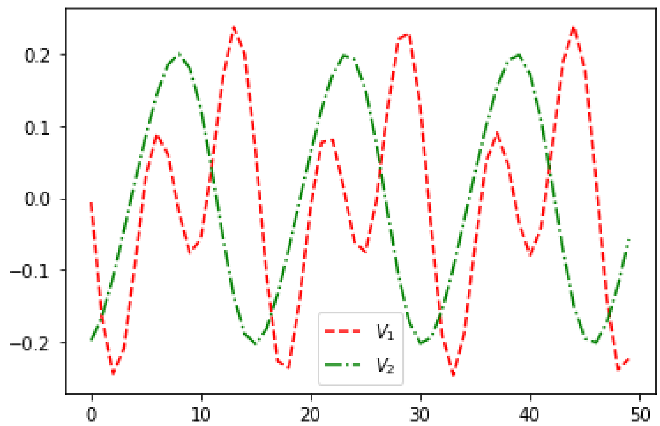
$$\frac{\sigma_1}{\sum\sigma_i} \approx 0.54$$

$$\frac{\sigma_1 + \sigma_2}{\sum\sigma_i} \approx 0.84$$

(d) Plot the first two columns of U and V in the SVD of matrix F obtained by calculat-ing $f(x,t)$ over a grid with 100 points in x and 50 points in t. Explain their meaning.

The columns of U are identified as spatial modes while components of V contain the time evolution of the modes. The singular value σ_i measure how much of the mode i contributes to the "energy" of the signal by f.





(4). To find a root of $f(x) = 0$, Newton's method tells to start with some initial guess x_0 and then to iterate following the scheme: $x_{n+1} = x_n - f(x_n)/f'(x_n)$.

(a) Use this method to find the root $x = 1$of $f(x) = x^2 - 1$.

$$f'(x) = 2x; => x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} = x_n - \frac{x_n^2 - 1}{2x_n} = x_n - \frac{1}{2}x_n + \frac{1}{2x_n} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right)$$

x_0	x_1	x_2	x_3	x_4
2	$\frac{5}{4}$	$\frac{41}{40}$	$\frac{3281}{3280}$	$\frac{21523361}{21523360}$

(b) hat is the range of initial conditions x_0 that give convergence to $x = 1$?
First, we can't divide by 0 => $x_0 \neq 0$;
Second, we can't divide by 0 => $x_n \neq 0$ let's find x_0 that leads to 0

$$x_{n+1} = 0 = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right)$$

$$0 = \left(x_n + \frac{1}{x_n}\right)$$

$$x_n = -\frac{1}{x_n}$$

$$x_n^2 = -1$$

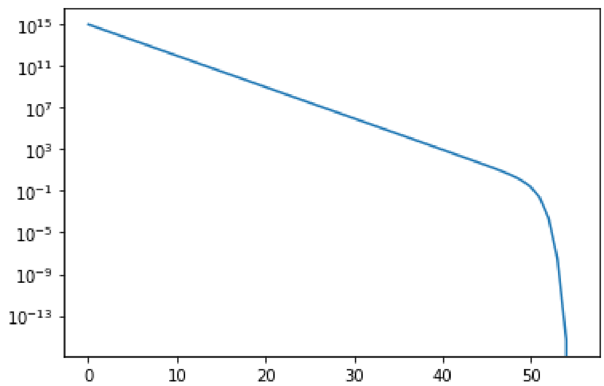
But x_n is real => no one x_0 leads to 0
Third, if $x_n \to x_{n+1}$; $x_{n+1} \to x_n$;

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right)$$

$$x_n = \frac{1}{2}\left(x_{n+1} + \frac{1}{x_{n+1}}\right)$$

After solving the system obtain:
$x_n = x_{n+1} = \pm 1$; - but it is the solution of $f(x) = 0$
Or we get solution for $x_n = x_{n+1}$ from complex space which is impossible.
Now we converge system to 1. But as you can see $x_{n+1} = \frac{1}{2}\left(x_n + \frac{1}{x_n}\right)$, so if $x_n > 0$ then $x_{n+1} > 0$ as sum of positive values. We want to $x_{n+1} = 1$ => $x_0 > 0$. We have only one positive point to converge ($x_{n+1} = 1$), that means our system will converge to it (can't converge to -1, because $-1 < 0$).
Finally, $x_0 \in (0, \infty)$.

(c) How fast do the iterations converge? Plot the error $e_n = |x_n - 1|$as a function of n(may be, in log scale).
Suppose there is typo in the task, because solution of $f(x) = 0$ is $x = \pm 1$, when $x = -1$ the error must be 0 but we get $e_n = |x_n - 1| = 2$
Let $e_n = |x_n| - 1$.
Go from 10^{15}on the chart:



Suppose:
$e_{k+1} = Ae_k^n$ divide by $e_{p+1} = Ae_p^n$

$$\frac{e_{k+1}}{e_{p+1}} = \frac{e_k^n}{e_p^n}$$

after taking log we get for n:

$$n = \frac{\log(e_{k+1}) - \log(e_{p+1})}{\log(e_k) - \log(e_p)}$$

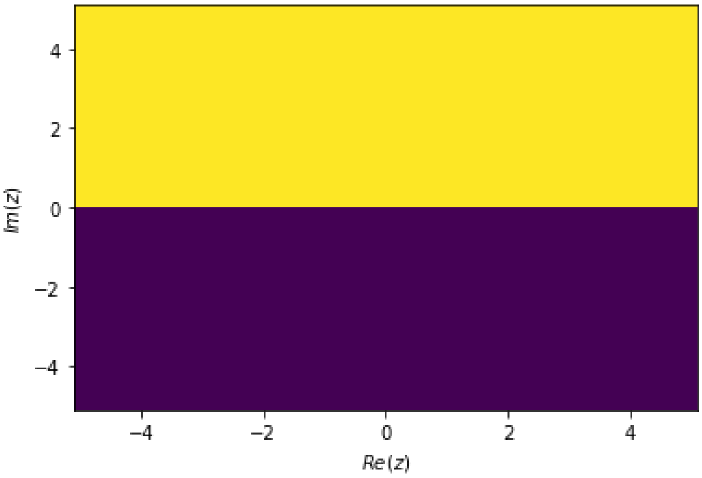
for near k and p I got $n \approx 1.9$ – the order of convergence. So, method converges very fast.

(5). Now apply the same Newton iterations as in the previous problem to the equation $f(x) = x^2 + 1 = 0$. Clearly, this equation has no real roots.

(a) The question is: What do the iterations do? Do they converge to anything?
$x_{n+1} = x_n - f(x_n)/f'(x_n)$. Where $f(x_n) \neq 0$ => x_{n+1}doesn't converge to anything.
$f(x) > 0 \ \forall\, x \in R$; $f'(x) = 2x > 0 \; if \; x > 0$; $f'(x) < 0 \; if \; x < 0$;
Though $-f(x_n)/f'(x_n)$ will lead x_{n+1} to 0.
It means that x_n aims to fall to $x = 0$, oscillate around 0.
What if $x_{n+1} = 0 = \frac{1}{2}\left(x_n - \frac{1}{x_n}\right)$ => $x_n = \pm 1$.
So, I think where is will oscillations, till x_n finally becomes ± 1 and next in iteration will division by 0, the end.

(b) How does the behavior of the iterations depend on the initial point x_0?
If it starts from $x_0 = 0$ it breaks. If it starts from x_0 leading to 0 it breaks ($x_0 = \pm 1$ and others which leads to x_0).
Otherwise x_n aims to fall to $x = 0$, oscillate around 0

(c) What if you start the iterations in the complex plane? Can you get convergence to the actual roots $\pm I$ of the equation? What are the domains of attraction of the roots?
Yes, we can get convergence to the actual roots $\pm I$ if start the iterations in the complex plane. $z_{n+1} = \frac{1}{2}\left(z_n - \frac{1}{z_n}\right)$.
Yellow is place of complex space where z_0 converges to I. Purple, where z_0 converges to $-I$ - domains of attraction of the roots.



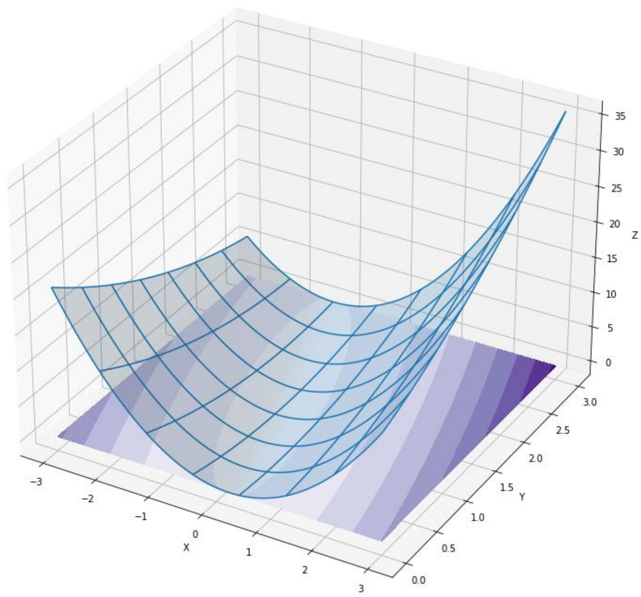
(6). Consider the function $f = 2x^2 + 2xy + y^2 - x - 2y$.

(a) Find its minimum analytically by representing f as $\frac{1}{2}u^TAu - b^Tu$. Plot the functiontogether with its contour levels using, for example, $surfc$ function in Matlab.

$f = \frac{1}{2}u^TAu - b^Tu$ where $A = A^T = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$; $b^T = [1 \quad 2]$; $u^T = [x \quad y]$

$$\nabla f = \frac{1}{2}u^T(2A) - b^T = u^TA - b^T = 0$$

$$u^T = b^TA^{-1} = \begin{bmatrix} -\frac{1}{2} & \frac{3}{2} \end{bmatrix}$$



(b) Now find the minimum using the gradient descent. Determine the step τ in the descent method.

Find τ_k such $\nabla f(x_k)^T \nabla f(x_{k+1}) = 0$.

NOTE: In my case I've found ∇f as a row, therefore $\nabla f(x_k)\nabla f(x_{k+1})^T = 0$.

$\nabla f(x_k)\left(u_{k+1}^T A - b^T\right)^{\mathrm{T}} = \nabla f(x_k)(Au_{k+1} - b) = 0$

$\nabla f(x_k)Au_{k+1} = \nabla f(x_k)b$

$\nabla f(x_k)A(u_k - \tau_k \nabla f(x_k)^T) = \nabla f(x_k)b$

$\nabla f(x_k)Au_k - \nabla f(x_k)A\tau_k \nabla f(x_k)^T = \nabla f(x_k)b$

$$\tau_k = \frac{\nabla f(x_k)Au_k - \nabla f(x_k)b}{\nabla f(x_k)A\nabla f(x_k)^T}$$

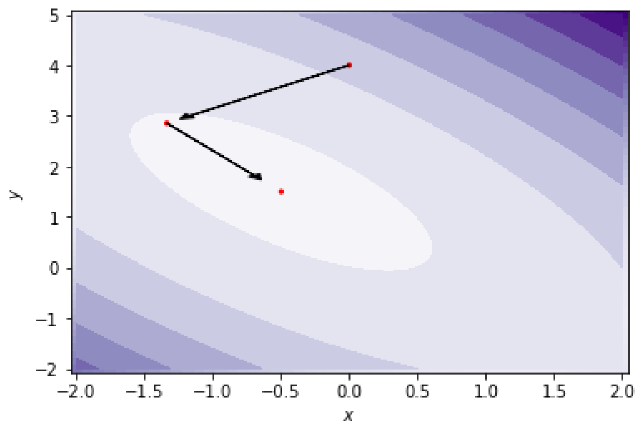
Using the algorithm $u_{k+1} = u_k - \tau_k \nabla f(x_k)^T$ found $u_{12}^T = [-0.5\,, 1.5]$, $u_0^T = [1{,}1]$; t_k almost always equals to 0.19 or 1.3.

(c) Starting with $(x_0, y_0) = (0{,}4)$, calculate the first two steps of the gradient descent explicitly and indicate on a single plot both the positions and the gradient vectors at those positions. Also plot the level curves off going through these points.

Let's $\tau_0 = 0.19$

$$u_1 = \begin{bmatrix}0\\4\end{bmatrix} - 0.19\begin{bmatrix}7\\6\end{bmatrix} = \begin{bmatrix}-1.34\\2.85\end{bmatrix}$$

$$u_2 = \begin{bmatrix}-1.34\\2.85\end{bmatrix} - 1.3\begin{bmatrix}-0.64\\1.03\end{bmatrix}$$



(d) Implement the descent algorithm in Matlab or Python and starting with the same initial condition as in (c) find the minimum within a tolerance of $tol = 10^{-6}$. How many iterations does it take to reach the minimum?

4 iterations to get: $x = -0.499999999339904$; $y = 1.500000000411$