

NATIONAL AND KAPODISTRIAN UNIVERSITY OF ATHENS

School of Science

Information Technologies in Medicine and Biology

Direction: *Bioinformatics*

Algorithms in Structural Bioinformatics

Postgraduate Student: *Begetis Nikolaos*

Professor: *Emiris Ioannis*

Deadline Date: 12/03/2013

Assignment 2

1. Structural alignment

In the first exercise of this assignment we have to compute the cRMSD of the x and y pointsets (Task (a)), then we have to compute the centroids and translate both sets to the origin, and using that, then, compute their new cRMSD (Task(b)). As a 3rd task (task (c)), we find the optimal rotation for alignment and afterwards, we apply it to one pointset and compute the optimal cRMSD. In task (d), for the input sets, we compute their distance d-RMSD, and by using that we can infer the final sets of task (c) without computing any d-RMSD. Subsequently, in task (e) we are asked if it is better to align set x with the z set, by repeating (a) to (c) steps, or by applying another transformation. Finally, in task (f) we repeat task (a)-(c) for x and y pointsets amplified by two eigenvectors.

The three sets-matrices are:

0	0	0
2	0	0
1	1	0
0	0	1

X

2	0	0
2	-2	0
3	-1	0
2	0	1

Y

0	0	0
4	0	0
2	2	0
0	0	2

Z

And the additions vectors for X and Y sets in task (f) are:

$X_5 = (-1, 1, -1)$ and $Y_5 = (0, 1, 1)$

So, to start with, for the task (a)-(f) we implemented a matlab function which computed and find for the task(a) the c-RMSD (Root Mean Square Deviation) as:

$$cRMSD = 2.2361$$

Then, for task(b) our implementation computes the centroid of line and columns values and translates them to common origin. Using that, the new c-RMSD is computed and it was found to be:

$$cRMSD_{new} = 1.3229$$

Afterwards, in task(c) we found the optimal rotation Q

$$Q = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

and using that we computed the optimal c-RMSD to be:

$$cRMSD_{opt} = 3.1524e^{-016} \approx 0$$

In the next step (task (d)), we were asked to compute the distance d-RMSD between the two pointsets and it was found to be:

$$dRMSD = 0$$

Now, it is worthy to mention that d-RMSD is not influenced from the rotation step, because the distance $\text{dist}(x_i, x_j)$ is equivalent to the distance $\text{dist}(Qx_i, Qx_j)$, $Q=1$. Moreover, d-RMSD is, also, not influenced from the second step, because no translation has an effect on d-RMSD, as $\text{dist}(x_i, x_j)$ is equivalent to $\text{dist}(x_i - x_c, x_j - x_c)$ and because of the absolute measure of the distance function the x_c distances are expunged. In general, the only ones that affect the distance dRMSD are scaling and distortion, which both do not occur in our sets. As a result, we can now conclude that we can guess the d-RMSD for the final sets in (c) without computing any d-RMSD, and it is 0, so in the next steps we will not have to compute dRMSD again.

For task (e), we can observe that set z is the same as set x multiplied by the value 2, so scaling is the function applied on x, which gives z. So, there is no need to repeat (a) to (c), as there exists no rotation. The only thing to be done is to find the centroids and translate both sets to the origin.

Finally, for the task (f) we changed the two pointsets x and y and we repeated the steps (a) – (c).

0	0	0
2	0	0
1	1	0
0	0	1
-1	1	-1

X

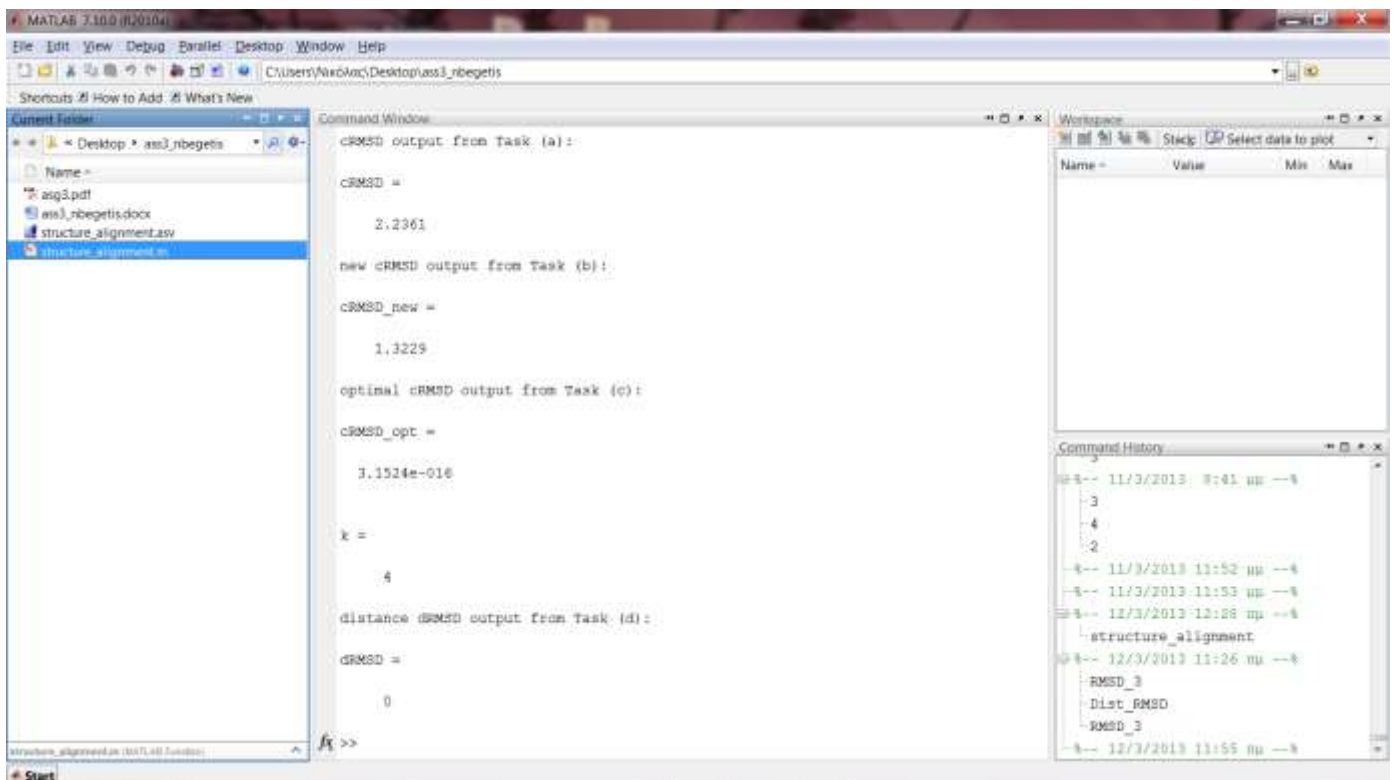
2	0	0
2	-2	0
3	-1	0
2	0	1
0	1	1

Y

The results we found were:

$$\begin{aligned}cRMSD &= 2.2361 \\ cRMSD_{new} &= 1.4967 \\ cRMSD_{opt} &= 0.8336\end{aligned}$$

More information can be found in code's comments and a screenshot of the results is collocated below:



The screenshot shows the MATLAB R2010a interface. The Command Window displays the following output:

```
cRMSD output from Task (a):  
cRMSD =  
2.2361  
  
new cRMSD output from Task (b):  
cRMSD_new =  
1.3229  
  
optimal cRMSD output from Task (c):  
cRMSD_opt =  
3.1524e-016  
  
k =  
4  
  
distance cRMSD output from Task (d):  
cRMSD =  
0
```

The Workspace window shows the following variables:

Name	Value	Min	Max
cRMSD	2.2361		
cRMSD_new	1.3229		
cRMSD_opt	3.1524e-016		
k	4		
cRMSD	0		

The Command History window shows the following commands:

```
1. 11/3/2013 8:41:11 pm -->  
2. 11/3/2013 11:52:11 pm -->  
3. 11/3/2013 11:53:11 pm -->  
4. 12/3/2013 12:28:11 pm -->  
5. 12/3/2013 11:26:11 pm -->  
6. 12/3/2013 11:55:11 pm -->
```

2. C-space

Unfortunately, an unexpected meeting occurred in my job and I did not find the time to complete this task. Nonetheless, if you ask me to make an implementation for this nice exercise email me.