

UNIVERSITY OF OSLO

Faculty of Mathematics and Natural Sciences

Exam:	IN5520 / IN9520 – Digital image analysis
Date:	Thursday December 10, 2020
Exam hours:	15.00-19.00 (4 hours)
Number of pages:	10 pages of sketches to a solution + 3 pages Appendix
Enclosures:	None

- Read the entire exercise text before you start solving the exercises.
Please check that the exam paper is complete.
If you lack information in the exam text or think that some information is missing, you may make your own assumptions, as long as they are not contradictory to the “spirit” of the exercise. In such a case, you should make clear what assumptions you have made.
- You should spend your time in such a manner that you get to answer all exercises.
If you get stuck on one question, move on to the next question.
- There are 8 exercises. They are weighted proportional to the number of sub-questions.
- Your answers should be **short**, typically a few sentences and / or a sketch should be sufficient.
- Do not give just a numerical answer, but demonstrate your reasoning.
- If you are using information from other sources than the textbook and the lecture notes, you should always include references to these external sources.
- During the exam, there will be a zoom-session for questions at 16.00. Ask questions orally. Look through the exercise set before this to identify all your questions.
- Lecture notes, slides, exercises, as well as any online information are allowed. However, it is FORBIDDEN to collaborate or communicate with any person about the exercises during the exam period. You may be subject to a control conversation regarding the authenticity of your exam submission.
Ref: <https://www.mn.uio.no/english/about/hse/corona/kontrollsamtale.html>.
Such a conversation will not impact grading of your exam, but it can lead to the institute raising a fraud case. Read more about fraud definitions here:
<https://www.uio.no/english/about/regulations/studies/studies-examinations/routines-cheating.html>

Good luck!

Exercise 1: Texture Analysis

Assume that you are given a gray level image of size $M \times N$ pixels with b bits per pixel.

- a) Describe how a normalized Gray Level Cooccurrence Matrix (GLCM) is computed, and which parameters this involves.

Answer: Textbook stuff, same as last year!

Eventual re-quantization of the input gray level image from $G = 2^b - 1$ to L gray levels should be motivated.

4 elements should be mentioned:

- 1. Initialize matrix of $G \times G$ (or $L \times L$) with 0's.*
- 2. Go through all $M \times N$ pixels where pixel pair of gray levels i and j a distance d pixels apart in direction θ are inside image, and add 1 at position (i, j) in GLCM.*
- 3. Finally, normalize by integer sum of matrix entries.*
- 4. Parameters: G (or eventually L), d , θ (or $\Delta x, \Delta y$).*

- b) Describe the procedure of extracting features from the matrix that directly describe the contents of the matrix, and thereby indirectly describes the texture being analysed; exemplified by a couple of well known features from two different feature categories.

Answer: Many of these GLCM features may be seen as a weighted sum of the cooccurrence matrix element values, where the weighting applied to each element is based on a given weighting function.

- 1. Weighting based on the value of the GLCM element, e.g., ASM or Entropy*
- 2. Weighting based on the position within the GLCM, IDM or Inertia.*

- c) Describe – without any equations - how a Gray Level Cooccurrence Matrix can be used for the purpose of globally thresholding a gray level image so that the resulting binary image has the smallest possible complexity, i.e. the smallest number of transitions between 0 and 1 in 4-connectivity.

Answer: Using $(\Delta x, \Delta y) = (0, 1)$ and $(1, 0)$, a nearest neighbor GLCM is obtained from the whole image. The GLCM in the opposite directions is the transpose of the GLCM. Adding the two gives us an isotropic GLCM in 4-connectivity.

Since we are mentioning “the smallest possible complexity, it may be tempting to propose a matrix feature that measures textural complexity. However, we are aiming for a global gray level threshold value that results in a minimal complexity in the output image. So, for a given threshold T on both axes of the GLCM, draw a vertical and a horizontal line through (T, T) . The total number of 0-1 and 1-0 transitions in the resulting binary image is the sum of entries in the two “off-diagonal rectangles” of the drawing. Complementary, the sum of entries in the “on-diagonal squares” is the sum of 0-0 or 1-1 steps in the resulting binary image, which do not contribute to the binary complexity.

So, finding the value of the threshold T that gives the smallest sum of the “off-diagonal rectangles” – or the largest sum of the “on-diagonal squares”, is the solution to the problem.

- d) How would you go about making the above approach to thresholding adaptive, i.e. obtaining a varying threshold surface over the image?

Discuss using both overlapping and non-overlapping windows.

Answer: By using a sliding window of a suitable size to obtain a GLCM and determining the complexity-optimal threshold value for the position of the center of the sliding window. Then move the window and repeat, until the whole image is covered. Eventually smooth the thresholding surface.

Moving the window one pixel at the time invites the use of GLCM updating, similar to the Huang, Yang, Tang algorithm for histogram updating in window-based median filtering.

(Ref: Huang, Yang, and Tang: “A fast two-dimensional median filtering algorithm”, IEEE TASSP 27(1), 13-18, 1979.)

If non-overlapping window positions are used, a sparse grid of threshold values is obtained, and an interpolation step is needed to complete the thresholding surface.

Exercise 2: Watershed Segmentation

- a) Explain the principal steps of Watershed Segmentation of graylevel images.

Answer:

- *Look at the graylevel image as a 3D topographic surface, $(x,y,intensity)$, with both local valleys and mountains.*
- *Assume that there is a hole at each minimum, and that the surface is immersed into a lake.*
- *The water will enter through the holes at the minima and flood the surface.*
- *To avoid that two different basins merge, a dam is built.*
- *Final step: the only thing visible would be the dams.*
- *The connected dam boundaries correspond to the watershed lines.*

- b) Explain how graylevel morphology can be used to detect local minima, prior to Watershed Segmentation.

Answer: A closing will fill in local minima of the graylevel image.

Closing = dilation followed by erosion, or $\min(\max(f))$

- c) Under what circumstances may this local minimum detection scheme become uncertain or even fail?

Answer: The size and shape of the structuring element (S) is important. If the local minimum is homogeneous over an area that is larger than S, the same minimum value may be detected in several connected pixel, and thus be uncertain. This will be of no consequence for the rest of the Watershed Segmentation. However, if the local minimum is flat over an extended area, no local minimum will be detected, and the segmentation scheme will fail.

This is in line with an aspect of filtering and template matching, that you need a filter/template that matches what you are intend to filter or detect.

- d) Given a binary object consisting of two partially overlapping squares, as shown in the right hand figure. Explain how you would split this object in two, using a distance transform and Watershed Segmentation.

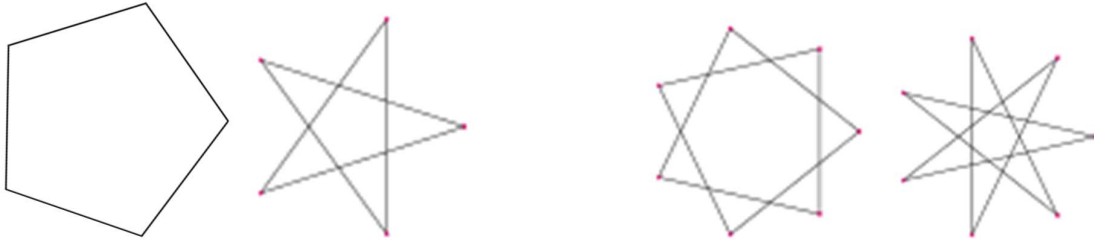


Answer: For each object pixel, compute the shortest distance to the background, giving a distance transform image. Then invert the distance image, and apply the watershed algorithm on this inverse distance image, giving a split line between the two concavity points on the contour of the object.

Exercise 3: Geometric Moments, Hough Transform and polygons

From an n-sided regular (equiangular and equilateral) polygon, a star-shaped string-object is produced by joining every m'th vertex of the polygon.

A common example is the *pentagram*, which has the same vertices as the pentagon, but connects every second vertex, giving a “stariness” of $\{5/2\}$, known as Schläfli ratio (figs to the left below).



- a) Describe the use of geometric moments and normal-representation Hough transform to detect this kind of non-overlapping star-shaped string-objects – limited to the $\{7/2\}$ and $\{7/3\}$ *heptagrams* (figures to the right above) - on a production line in the domestic decorations industry, based on digital images of one object per image, and sorting the objects by their “stariness” or Schläfli ratio.

Answer: We use first order moments to find the center of mass (CM), and move the origin to CM.

$$m_{10} = \sum_x \sum_y x b(x, y) = \bar{x} m_{00} \Rightarrow \bar{x} = \frac{m_{10}}{m_{00}}$$

$$m_{01} = \sum_x \sum_y y b(x, y) = \bar{y} m_{00} \Rightarrow \bar{y} = \frac{m_{01}}{m_{00}}$$

Normally, the input to HT is a thresholded gradient magnitude image. Here, the image of the string-object is equivalent to a binary gradient magnitude image.

For a given heptagram, the Hough transform (HT) will give 7 peaks in the HT-parameter domain (ρ, θ) , all at the same ρ -value, and at equidistant θ -values. The ρ -value will be larger for the $\{7/2\}$ heptagram compared to the $\{7/3\}$ heptagram.

Note that the corresponding 7 lines will be longer than the actual sides of the heptagram, but we do not need to establish start- and endpoints. To establish the Schläfli ratio, we need to establish the vertices, and using the formalism of slide 45 of lecture 3 on HT, we can solve for every vertex, and then check how many vertices there are (7) and how many are passed ($m = 2$ or 3), as we follow a line from one side of the image, through one vertex to the next vertex, and thus establish whether the heptagram is $\{m/n\} = \{7/2\}$ or $\{7/3\}$.

- b) Describe the use of convex hull (CH) based descriptors as a proxy for the above “stariness” measure of the two *heptagrams* above.

Answer: The convex hull CH is the smallest convex set S that contains the heptagram, and the set difference $H - S$ is called the convex deficiency D.

Useful descriptors that have been mentioned are

1. CH “solidity” aka “convexity” = (object area)/(CH area)
2. Number of components of convex deficiency and distribution of component areas
3. Relative location of points where we go in and out of the convex deficiency
4. Points of local maximal distance to CH.

The convex hull area is the same for the two heptagrams, as they are constructed from the heptagon.

But the value of S is smaller for the $\{7/3\}$ heptagram => solidity is smaller for smaller for $\{7/3\}$.

2, 3 & 4 are obviously the same for the two heptagrams.

- c) Describe other geometric features that would be useful to distinguish between “stariness” of different m -values for a regular star polygon having n vertices exemplified by the two heptagrams above.

Answer: Signature, i.e., distance from center of mass versus angle, along the outer boundary of the star polygon will give us the number of vertices of the polygon (n). And the minimum distances from the center of mass will distinguish between different “stariness”, since a line connecting every m 'th vertex of a polygon will pass closer to the center of mass as m is increased from $m=2$ to $m=3$.

Compactness (very simple measure) $\gamma = P^2 / (4\pi A)$, where P = Perimeter, A = Area, For a circular disc, γ is minimum and equals 1. Compactness attains high value for complex object shapes, Here, P is larger and A is larger for the $\{7/2\}$ heptagram than for the $\{7/3\}$ heptagram.

Irregularity, either defined as the ratio of the centered enclosing circle to the area A , or the ratio of the radii of the maximum and minimum centered circles, will also distinguish between the two heptagrams.

Exercise 4: For PhD-stud.: Feature extraction for a classification task

- a) You are given the task of designing a system for fish species classification. Based on fish caught from a ship you should classify them as either salmon or cod. Two example images are shown below (left: salmon, right:cod).



Discuss briefly three image-based features that you would suggest for a binary classification problem of fish as either salmon or cod. Does your suggested feature require that any particular camera system is used? Consider if the features should characterize subparts of the fish. Does your suggested algorithm require any preprocessing?

Answer:

Any good discussion will give high score!

A RGB image is better than a grayscale image. Color is one feature class, texture is another, in addition to larger brighter parts on the lower half of salmon. The location of fins is also different, and «beard» is unique for cod. Segmentation for localization of the fish and normalization of contrast/RGB-HSV-conversion might be necessary.

- b) Your first dataset consisted of 100 cod and 15 salmon.
How would you give the accuracy of your classifier?

Answer: A confusion matrix where you give the accuracy per class is e.g. better than just the number of correctly classified fish over all classes.

Exercise 5: Support vector machine classification

- a) Explain why SVM classifiers often focus on the sign of a dot product.

Answer: The dot product between a sample x and a point on the decision boundary measures the cosine of the angle between the boundary and the vector from x to a point on the boundary. To decide the class for a binary classifier, we only need to know on which side of the boundary we are.

To get full score here, we ask you to state which lines the dot product relate to. Note that it is different to when using a Gaussian classifier (where the line connecting the class means play a role).

- b) Explain by your own words what ξ_i is in the equation

$$J(w, w_0, \xi) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N I(\xi_i)$$

$$\text{where } I(\xi_i) = \begin{cases} 1 & \xi_i > 0 \\ 0 & \xi_i = 0 \end{cases}$$

and ξ is the vector of parameters ξ_i .

Answer: They denote slack variables, and are points that are between the margins, corresponding either to points which are inside the margin, but correctly classified ($0 < \xi < 1$) or points which are wrongly classified ($\xi > 1$). Point correctly classified outside the margin have $\xi = 0$.

- c) Describe briefly how a set of binary SVM-classifiers can be used on a multiclass problem.

Answer: Train pairwise classifiers on all class pairs. To decide the class for a new sample, evaluate the class label based on all pairwise classification, and choose majority. This can create ties.

We also state in the lecture note that you can train one-vs-all, but in that case you need the size of the output, not just the sign, to decide between all one-vs-all possibilities.

- a) Consider the figure below for an example with 3 classes. The figure shows the SVM boundaries for 3 class pairs. The blue/black line is the boundary between class 1 and 2, the red line the boundary between class 1 and 3, and the green line the boundary between class 2 and 3. All point within a region will be classified as the same class. For each pairwise classifier, the signs denote the classes, - is the class with the smallest label number and + the class with the highest label number of the pair.

Based on your candidate number, select a point in region 1-5 and describe how it is classified.

Let d be the last digit in your candidate number.

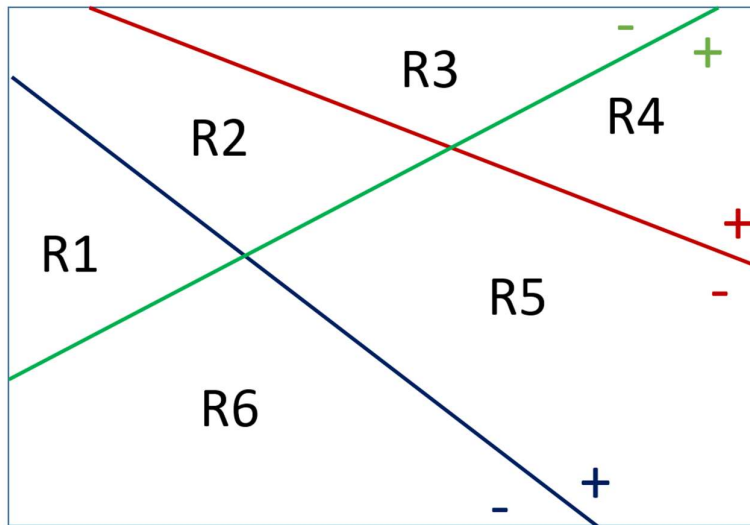
If $d=0-1$: select a point in region 1

If $d=2-3$: select a point in region 2

If $d=4-5$: select a point in region 3

If $d=6-7$: select a point in region 4

If $d=8-9$: select a point in region 5



Answer: The regions will have the following binary class prediction:

R1: Class 1 vs 2: 1

Class 1 vs 3: 1

Class 2 vs 3: 2

Most frequent value: 1

R2: Class 1 vs 2: 2

Class 1 vs 3: 1

Class 2 vs 3: 2

Most frequent value: 2

R3: Class 1 vs 2: 2

Class 1 vs 3: 3

Class 2 vs 3: 2

Most frequent value: 2

R4: Class 1 vs 2: 2

Class 1 vs 3: 3

Class 2 vs 3: 3

Most frequent value: 3

R5: Class 1 vs 2: 2

Class 1 vs 3: 1

Class 2 vs 3: 3

Most frequent value: 2

R6: Class 1 vs 2: 1

Class 1 vs 3: 1

Class 2 vs 3: 3

Most frequent value: 1

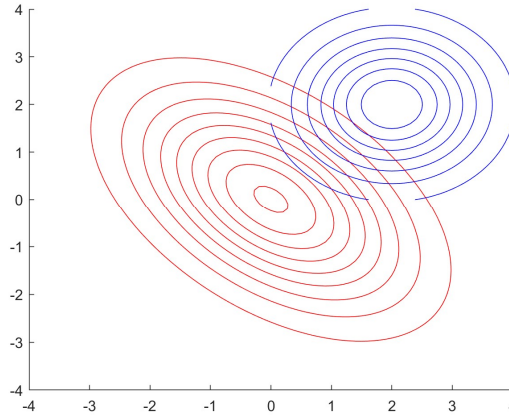
Exercise 6: A Gaussian classifier

In this exercise we will work with a Gaussian classifier and two classes defined with:

$$\mu_1 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \quad \Sigma_1 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix}, \quad \Sigma_1^{-1} = \begin{bmatrix} 0.667 & 0.333 \\ 0.333 & 0.667 \end{bmatrix}, \quad |\Sigma_1| = 3$$

$$\mu_2 = \begin{bmatrix} 2 \\ 2 \end{bmatrix}, \quad \Sigma_2 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \Sigma_2^{-1} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad |\Sigma_2| = 1$$

The classes are outlined in the plot below:



We are also given a set of samples that we will classify. There are 10 samples in the table. Select the sample that corresponds to the last digit in you candidate number

Candidate number (last digit)	Test sample	Candidate number (last digit)	Test sample
0	(1,1)	5	(1,2)
1	(3,0)	6	(0,3)
2	(0,1)	7	(2,0)
3	(2,1)	8	(0,2)
4	(1,0)	9	(3,1)

Compute the probability that your sample x belongs to class 1 and class 2, $p(x|\omega=1)$ and $p(x|\omega=2)$. Include all your computations here.

Answers:

Note that there was a typo in the inverse covariance matrix for class 2, which should also be diagonal. Both answers with the listed and the corrected matrix are given full score.

The formula is: (no prior probabilities given, so we assume them equal)

$$P(\omega_j|x) = p(x|\omega_j)P(\omega_j) = \frac{1}{(2\pi)^{d/2}|\Sigma_j|^{1/2}} \exp\left[-\frac{1}{2}(x - \mu_j)^t \Sigma_j^{-1}(x - \mu_j)\right] P(\omega_j)$$

Note that we ask for a probability, so if you only compute your answer based on a log-discriminant function this does not represent probabilities. Full score are given if you compute probabilities with or without the

prior probability (as long as you state what you do). Based on your computations, you get a score even if the answer is slightly wrong.

Sample	$p(x class1)$	$p(x class2)$	Sample	$p(x class1)$	$p(x class2)$
0	0.0169	0.0293	5	0.0045	0.0483
1	0.0023	0.0065	6	0.0023	0.0065
2	0.0329	0.0065	7	0.0121	0.0108
3	0.0045	0.0483	8	0.0121	0.0108
4	0.0329	0.0065	9	0.0045	0.000

Exercise 7: Subsets of features

In this exercise, we are working with a dataset with 4 features. The mean and covariance matrix for all the samples are:

$$\mu = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \text{ and } \Sigma = \begin{bmatrix} 2 & -1 & 0.5 & 0 \\ -1 & 3 & -1 & -1 \\ 0.5 & -1 & 3 & -1 \\ 0 & -1 & -1 & 2 \end{bmatrix}$$

The eigenvalues and eigenvectors of this matrix are:

$$\lambda_1 = 0.53 \quad v_1 = [-0.21 \quad -0.54 \quad -0.44 \quad -0.67]^T$$

$$\lambda_2 = 1.56 \quad v_2 = [-0.88 \quad -0.22 \quad 0.32 \quad 0.24]^T$$

$$\lambda_3 = 3.43 \quad v_3 = [0.07 \quad -0.40 \quad -0.59 \quad 0.58]^T$$

$$\lambda_4 = 4.47 \quad v_4 = [0.40 \quad -0.70 \quad 0.58 \quad 0.05]^T$$

- a) Now we do feature selection. Select feature k and m depending on the last digit d in your candidate number:
- $d=0-1$: $k=1, m=2$
 - $d=2-3$: $k=1, m=3$
 - $d=4-5$: $k=1, m=4$
 - $d=6-7$: $k=2, m=3$
 - $d=8-9$: $k=2, m=4$

Set up the covariance matrix for your selected features.

See appendix for solutions for all candidate numbers.

Answer: For feature 1,3 it is:

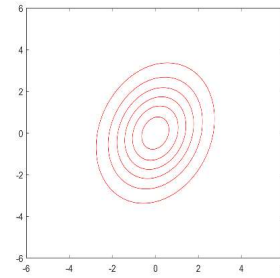
$$\Sigma = \begin{bmatrix} 2 & 0.5 \\ 0.5 & 3 \end{bmatrix}$$

- b) We now want to visualize this covariance matrix as an ellipse. Set up an expression for computing the minor and major axis for this ellipse, and compute the values for your selected combination of 2 features.

*Answer: From lecture notes on multivariate classification, slide 44, we find the formula
It is based on the eigenvalues and eigenvectors of the 2x2 covariance matrix in b)*

$$\lambda_1 = 3.20 \quad v_1 = [-0.92 \quad 0.38]^T$$

$$\lambda_2 = 1.79 \quad v_2 = [0.38 \quad 0.92]^T$$



The lecture notes state that the major axis of the ellipse is $2\sqrt{\lambda_1}v_1$ and the minor axis is $2\sqrt{\lambda_2}v_2$

For the selected values we get: $2\sqrt{\lambda_1}v_1 = [1.37 \ 3.30]^T$, $2\sqrt{\lambda_2}v_2 = [-2.47 \ 1.07]^T$

Comment: either computed values or correct equation gives full score.

Comment: the course also gives you other measures to estimate the direction of a matrix, e.g. using second-order moments. Solutions based on this are also given full score.

Exercise 8: Statistics

A city has two hospitals. In the biggest hospital, about 50 children are born each day, in the smaller hospital about 20. It is common knowledge that about 50% of all newborn are girls, but the exact percentage in any given hospital may vary from day to day. Sometimes it may be lower than 50%, sometimes higher.

For a period of time, each hospital recorded the number of days on which more than 60% of the babies born were girls. Which hospital do you think recorded more such days?

- 1) The larger hospital
- 2) The smaller hospital
- 3) About the same (i.e., within 5% of each other)

Please explain the reason(s) for your answer, and try to relate to another well known classification problem!

Answer: The correct answer is alternative 2) The smaller hospital, simply because a small sample is more likely to deviate from the 50/50 girl/boy ratio, just as you risk selecting nonsensical classification features if you have a small sample of images.

This is an example from the bestselling book “Thinking Fast and Slow” (Penguin Books 2011) by the Nobel laureate Daniel Kahneman (Appendix A, p 420-21): “53 of 95 undergraduates (56%) chose the last alternative, judging the probability of obtaining more than 60% girls to be the same in the small and the large hospital, presumably because these events are described by the same statistic and are therefore equally representative of the general population. In contrast, sampling theory entails that the expected number of days on which more than 60% of the babies are girls is much greater in the small hospital than in the large one, simply because a large sample is less likely to stray from 50%. This fundamental notion of statistics is evidently not part of people’s repertoire of intuitions.” Only 21 of 95 undergraduates (22%) chose the correct second alternative.

So how will the IN5520/9520 students be distributed?

And is it at all possible for us to compare our MSc and PhD students, given the small sample size? Since our students have been confronted with the “curse of dimensionality” in small sample size situations, we expect them to do better than Kahneman’s undergraduates! But again, do we have a sufficient sample size to compare with his 95 students?

In retrospect, our students did much better than Khaneman’s undergraduates; everybody answered “the smaller hospital”. But 4 out of 27 stated that the *reason* was that it took a smaller number of girls to get to 60% in the smaller hospital, ignoring that a larger number of babies were to be born in the larger hospital anyway. The reason is simply that a larger sample (in the larger hospital) is less likely to deviate from a 50/50 distribution.

Thank You for Your Attention!

Appendix – Solution to exercise 7 for all candidate numbers

Covariance matrix: $m_{12} =$

-1 3
Eigenvalues: $d_{12} =$
1.3820 0
0 3.6180

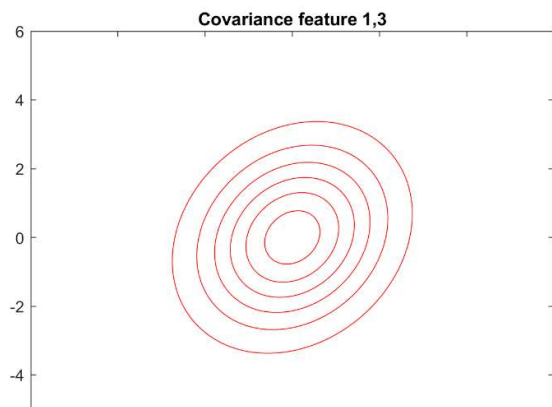
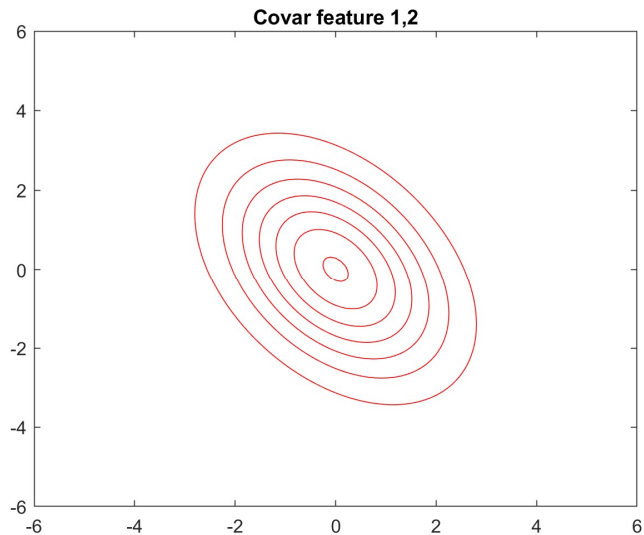
Eigenvectors: $v_{12} =$
-0.8507 -0.5257
-0.5257 0.8507

Major and minor axis of ellipse:
 $ax_{12_max} =$
-2.0000
3.2361

$ax_{12_min} =$
-2.0000
-1.2361

$m_{13} =$
2.0000 0.5000
0.5000 3.0000
Eigenvalues: $d_{13} =$
1.7929 0

Note: we do not ask for the covar sketch!



0 3.2071

Eigenvectors: v13 =

-0.9239 0.3827
0.3827 0.9239

ax13_max =

1.3706
3.3090

ax13_min =

-2.4741
1.0248

m14 =

2 0
0 2

d14 =

2 0
0 2

v14 =

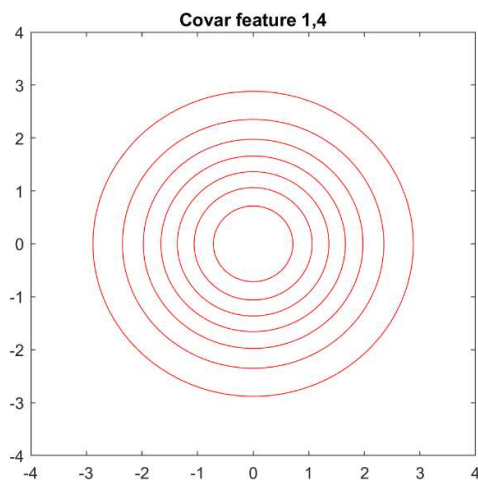
1 0
0 1

ax14_max =

0
2.8284

ax14_min =

2.8284
0



$m_{23} =$

$$\begin{matrix} 3 & -1 \\ -1 & 3 \end{matrix}$$

$d_{23} =$

$$\begin{matrix} 2 & 0 \\ 0 & 4 \end{matrix}$$

$v_{23} =$

$$\begin{matrix} -0.7071 & -0.7071 \\ -0.7071 & 0.7071 \end{matrix}$$

$ax_{23_max} =$

$$\begin{matrix} -2.8284 \\ 2.8284 \end{matrix}$$

$ax_{23_min} =$

$$\begin{matrix} -2 \\ -2 \end{matrix}$$

$m_{24} =$

$$\begin{matrix} 3 & -1 \\ -1 & 2 \end{matrix}$$

$d_{24} =$

$$\begin{matrix} 1.3820 & 0 \\ 0 & 3.6180 \end{matrix}$$

$v_{24} =$

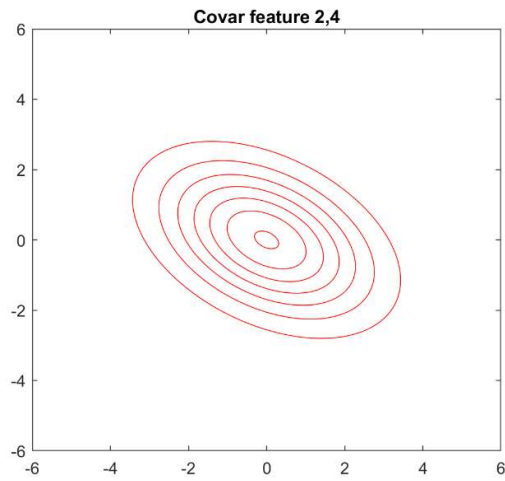
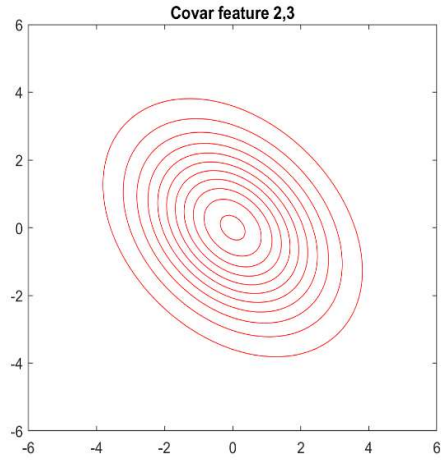
$$\begin{matrix} -0.5257 & -0.8507 \\ -0.8507 & 0.5257 \end{matrix}$$

$ax_{24_max} =$

$$\begin{matrix} -3.2361 \\ 2.0000 \end{matrix}$$

$ax_{24_min} =$

$$\begin{matrix} -1.2361 \\ -2.0000 \end{matrix}$$



m34 =

3 -1
-1 2

d34 =

1.3820 0
0 3.6180

v34 =

-0.5257 -0.8507
-0.8507 0.5257

ax34_max =

-3.2361
2.0000

ax34_min =

-1.2361
-2.0000

