

## NPFL097 Warm-up test, Feb 25, 2020

1. Define convex function.
2. What is gradient?
3. Find a growing function which maps  $R$  to  $(-\infty, \infty)$ .
4. Given joint distribution  $p(A, B)$ , what are the terms for  $p(A)$  and  $p(A|B)$ ?
5. Explain “curse of dimensionality”.
6. Explain the main difference between the frequentist and Bayesian interpretation of probability.
7. Derive Bayes’ theorem.
8. Define independence (independent random variables  $X$  and  $Y$ ).
9. Define conditional independence (variable  $X$  independent of  $Y$  given  $Z$ ).
10. What does it mean that a collection of random variables is i.i.d. (independent and identically distributed)?
11. What is the relation between a probability density function and associated cumulative distribution function?
12. What is correlation?
13. What is variance?
14. What is covariance matrix?
15. Let’s suppose that the sequence (1, 3, 4, 4, 8) is drawn from  $\mathcal{N}(\mu, \sigma^2)$ . What are the values of  $\mu$  and  $\sigma^2$  (according to Maximum Likelihood)?
16. What can you say about a multidimensional Gaussian with covariance matrix having zeros everywhere outside its diagonal?
17. Which types of random variables cannot be modeled by Gaussian distributions?
18. Explain the difference between generative models and discriminative models.
19. What are Monte Carlo methods?
20. How can you generate samples from a uniform distribution in 3D unit ball?
21. Explain the difference between classification and regression.
22. What is separation boundary?
23. What does it mean that two sets of points are linearly separable.
24. Plot the sigmoid function used in logistic regression ( $f(z) = \frac{1}{1+e^{-z}}$ ).
25. Design feature transformation functions which make the following sets linearly separable: (a)  $A = (2, 0)$  and  $B = (1, 0), (3, 0), (2, 1), (2, -1)$ , (b)  $A = (0, 0), (2, 2)$  and  $B = (1, 1), (3, 3)$ .
26. Explain how k-means clustering works.
27. Illustrate underfit/overfit problems in regression by fitting a polynomial function to a sequence of points in 2D.
28. Illustrate underfit/overfit problems in classification by modeling two (partially overlapping) classes of points in 2D.
29. Plot a typical dependence of training and test error rates (vertical axis) on training data size (horizontal axis).
30. Plot a typical dependence of training and test error rates (vertical axis) on model complexity (horizontal axis).
31. What problem is typically signalled by test error being much higher than training error?
32. What problem is typically signalled by test and training errors being stabilized after a limited portion of training data?
33. What is regularization used for?
34. Plot the following functions: (a)  $f(x) = x \exp(x)$ , (b)  $f(x) = \ln \frac{x^2-1}{x^2+1}$ , (c)  $f(x) = \left| \frac{x-1}{1-2x} \right|$ , (d)  $f(x) = \sqrt{1 - \exp(-x^2)}$