Maschinelles Lernen 1:
1. Klausur WS18/19

4 multiple choice fragen: 20 Punkte 4 x 5
nur eine richtige antwort

What …. is not a discriminant function:
a) $P(w_c|x)$
b)$P(x|w_c)*P(w_c)$
c)$P(w_c|x)*P(w_c)^{-1}$
d)$P(x|w_c)^2*P(w_c)^2$

What is likely an overfitted estimator?
a) High bias model
b) High variance model
c) Low bias model
d) Low variance model

What does Fisher discriminant optimize?
a)..
b)..
c) Maximize ratio Within class variance to between class variance
d) Minimize ratio Within class variance to between class variance

What does the constant C stand for in SWM?
a) ability of the decision boundary to be out of the margin
b) number of points not being classified correctly
c)
d)

Boosting: 15 Punkte
Genau wie in der Hausaufgabe, Gewichte der Punkte und gewichte der Classifier angeben

Kernel: 20 Punkte 4 x 5
Genau wie in Hausaufgabe:
a) Zeige das $k_3 = \alpha k_1 + \beta k_2$ wieder ein positive semidefiniter Kernel ist.
Alpha, beta $>= 0$

b) Zeige das wenn alpha oder beta < 0 , dass dann die obere bedingung nicht mehr stimmt.

c) Find a mapping for k_3 asssuming that $k_2 = \langle \Phi_2(x_i),\Phi_2(x_j)\rangle$,
$k_1 = \langle \Phi_1(x_i),\Phi_1(x_j)\rangle$ and prove that it satisfies $k_3 = \langle \Phi_3(x_i),\Phi_3(x_j)\rangle$

d)
Do the same as in c but for $k4(x_i,x_j)=k3(x_i,x_i)k3(x_j,x_j)$

Kernel: 20 Punkte (4 × 5)

Genau wie in Hausaufgabe:

a)

Zeige, dass

$$k_3 = \alpha k_1 + \beta k_2$$

wieder ein **positiv semidefiniter Kernel** ist.
Voraussetzung: $\alpha, \beta \geq 0$.

$$\sum_i \sum_j c_i c_j \left(\alpha k(x_i, x_j) + \beta k(x_i, x_j) \right)$$

b)

Zeige, dass wenn α oder $\beta < 0$, dann die obere Bedingung nicht mehr stimmt.

c)

Finde eine **Mapping-Funktion** für k_3, unter der Annahme:

$$k_2 = \langle \Phi_2(x_i), \Phi_2(x_j) \rangle, \quad k_1 = \langle \Phi_1(x_i), \Phi_1(x_j) \rangle$$

und beweise, dass dies die Bedingung

$$k_3 = \langle \Phi_3(x_i), \Phi_3(x_j) \rangle$$

erfüllt.

$$k_3 = \alpha \langle \Phi_1(x_i), \Phi_1(x_j) \rangle + \beta \langle \Phi_2(x_i), \Phi_2(x_j) \rangle$$

$$= \alpha \Phi_1^T(x_i) \Phi_1(x_j) + \beta \Phi_2^T(x_i) \Phi_2(x_i)$$

$$= \begin{bmatrix} \sqrt{\alpha}\Phi_1^T(x_1) & \sqrt{\beta}\Phi_2^T(x_1) \end{bmatrix} \begin{bmatrix} \sqrt{\alpha}\Phi_1(x_1) \\ \sqrt{\beta}\Phi_2(x_2) \end{bmatrix}$$

$$= \left\langle \begin{bmatrix} \sqrt{\alpha}\Phi_1(x_1) \\ \sqrt{\beta}\Phi_2(x_2) \end{bmatrix}, \begin{bmatrix} \sqrt{\alpha}\Phi_1(x_1) \\ \sqrt{\beta}\Phi_2(x_2) \end{bmatrix} \right\rangle$$

d)

Mache dasselbe wie in **c**, aber für

$$k_4(x_i, x_j) = k_3(x_i, x_i) k_3(x_j, x_j)$$

$$= \phi_3(x_1)^T \phi_3(x_1) \; \phi_3(x_2)^T \phi_3(x_2)$$

$$= \underbrace{\left(\phi_3(x_1)^T \phi_3(x_1) \right)^T}_{\phi_4} \phi_3(x_2)^T \phi_3(x_2)$$

Lagrang: 25 Punkte 5 x 5

$$\frac{\partial L}{\partial w} = 2w - 2\lambda\Sigma^{-1}w = 0$$
$$w = \lambda\Sigma^{-1}w$$
$$\Sigma w = \lambda w$$

Sigma is the covariance matrix of the dataset
max w w^T s.t. w Sigma^{-1} $w^T = 1$

a) Derive Lagrangian $L = ww^T - \lambda(w\Sigma^{-1}w^T - 1)$

$$\max ww^T = \lambda\Sigma^{-1}ww^T = \lambda$$

b) Show that the problem is a Eigenvalue problem of Sigma.

c) Show that the solution is the eigenvector corresponding to the largest eigenvalue.

d) Derive a closed form solution for x^t given:

$$L = zx_{t-1} - \lambda(z\Sigma^{-1}z^T - 1)$$
$$\frac{\partial L}{\partial z} = x_{t-1} - 2\lambda\Sigma^{-1}z^T = 0 \qquad x_{(t-1)} = 2\lambda\Sigma^{-1}z^T$$

max z x^{t-1} s.t z Sigma^{-1} z^T = 1

$$\max 2\lambda z\Sigma^{-1}z^T = 2\lambda$$

e) What algorith does this look similar to regarding PCA mentioned in ML1?

Regression: 20 Punkte 5 , 15

Klassisches Regressionsproblem:

min Summe (y_i – wx_i +b)² oder so.

a) Zeigen das min w^TXXw – 2yXw das gleiche problem ist.

b) Erstelle die Matrixen wie in der hausaufgabe für diesen qudratic solver. Q,b,A,l
Da war auch noch eine Konstante C gegeben.