Aufgabe 06.1

$$\min_{x \in \mathbb{R}^n} \| Ax - y \|_2^2 \qquad P \in \mathbb{R}^{(n+k)\times(n+k)}, \; p \in \mathbb{R}^{n+k} \Rightarrow P\begin{bmatrix} x \\ \lambda \end{bmatrix} = p$$

s.t. $Bx = b$

① $\min_{x \in \mathbb{R}^n} \|Ax - y\|_2^2 = (Ax-y)^T(Ax-y) = y^Ty - 2y^TAx + x^TA^TAx.$

The Lagrangian is giving by:

$$L(x, \lambda) = \cancel{\tfrac{1}{2}} \|Ax - y\|_2^2 + 2\lambda'(Bx-b)$$

$$= y^Ty - 2y^TAx + x^TA^TAx + 2\lambda^T(Bx-b)$$

② Dann
$$\begin{cases} \nabla_x L = -2y^TA - 2B^T\lambda + 2A^TAx = 0 \\ \quad \Rightarrow A^TAx = A^Ty - B^T\lambda \\ \nabla_\lambda L = 2(Bx-b) = 0 \\ \quad \Rightarrow Bx = b \end{cases}$$

③ D.h haben wir $\begin{bmatrix} A^TA & B^T \\ B & 0 \end{bmatrix} \cdot \begin{bmatrix} x \\ \lambda \end{bmatrix} = \begin{bmatrix} A^Ty \\ b \end{bmatrix}$

(mit KKT-condition)

Aufgabe 08.2

$$\min_{w,b,\varepsilon,\rho} \ \frac{1}{2}\|w\|^2 - \nu\rho + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i \qquad 1)$$
$$\text{s.t.} \quad y_i(w^T x_i + b) \geq \rho - \varepsilon_i \quad \text{for all } i$$
$$\varepsilon_i \geq 0 \ , \ \rho \geq 0$$

① $$L(w,\varepsilon,b,\rho,\alpha,\beta,\delta) = \frac{1}{2}\|w\|^2 - \nu\rho + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i \qquad 2)$$
$$- \sum_{i=1}^{n}\left(\alpha_i\left(y_i(\langle x_i, w\rangle + b) - \rho + \varepsilon_i\right) + \beta_i\varepsilon_i - \delta\rho\right)$$

Then min w,ε,b,ρ and max. α,β,δ $\Rightarrow \max_{\alpha}\min_{w,b} L$.

mit $\frac{\partial}{\partial b}L(w,b,\alpha)=0$, $\frac{\partial}{\partial w}L(w,b,\alpha)=0$ 3)

② haben wir $\Rightarrow w = \sum_{i=1}^{m}\alpha_i y_i x_i$, $\alpha_i + \beta_i = 1/n$, $\sum_{i=1}^{n}\alpha_i y_i = 0$, $\sum_{i=1}^{n}\alpha_i - \delta = \nu$ (4)

③ Dann 4) in 2) haben wir

$$\underset{\alpha\in\mathbb{R}^m}{\text{maximize}} \quad W(\alpha) = -\frac{1}{2}\sum_{i,j=1}^{m}\alpha_i\alpha_j y_i y_j k(x_i,x_j) \qquad (k(x_i,x_j) = e^{-\gamma\|x_i-x_j\|^2})$$
$$\text{s.t.} \quad 0 \leq \alpha_i \leq \frac{1}{n}$$
$$\sum_{i=1}^{m}\alpha_i y_i = 0$$
$$\sum_{i=1}^{m}\alpha_i \geq \nu$$

$$\Leftrightarrow \ \cancel{f} \quad \frac{1}{2}\|w\|^2 + C\left(-\nu\rho + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i\right)$$

Ex. in ② $\nabla_w L(w,b,\alpha) = w - \sum_{i=1}^{n}\alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^{n}\alpha_i y_i x_i$

$\nabla_b L(w,b,\alpha) = -\sum_{i=1}^{n}\alpha_i y_i = 0 \Rightarrow \sum_{i=1}^{n}\alpha_i y_i = 0$

$\nabla_{\varepsilon_i} L(w,b,\alpha) = \beta_i + \alpha_i - \frac{1}{n} = 0 \Rightarrow \alpha_i + \beta_i = \frac{1}{n}$

$\nabla_\rho L(w,b,\alpha) = \sum_{i=1}^{n}\alpha_i - \delta - \nu = 0 \Rightarrow \sum_{i=1}^{n}\alpha_i - \delta = \nu$

Ex in ③ $L = \frac{1}{2}\|w\|^2 - \nu\rho + \frac{1}{n}\sum_{i=1}^{n}\varepsilon_i - \sum_{i=1}^{n}\left(\alpha_i\left(y_i(\langle x_i, w\rangle + b) - \rho + \varepsilon_i\right) + \beta_i\varepsilon_i - \delta\rho\right)$

$= \frac{1}{2}\sum_{i=1}^{n}\alpha_i y_i x_i^T\sum_{j=1}^{n}\alpha_j y_j x_j - \sum_{i=1}^{n}\alpha_i y_i x_i^T\sum_{j=1}^{n}\alpha_j y_j x_j - b\sum_{i=1}^{n}\alpha_i y_i + \sum_{i=1}^{n}\alpha_i$

$= -\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}\alpha_i\alpha_j y_i y_j x_i^T x_j$

mit $\sum_{i=1}^{m}\alpha_i y_i = 0$, $\sum_{i=1}^{m}\alpha_j \geq \nu$.

Aufgabe 06.3

$\min_{w,b} \frac{1}{2}\|w\|^2$

s.t $y_i(w^T x_i + b) \geq 0$ for i

1) jetzt $L(w,b,\alpha) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{N} \alpha_i (y_i (w^T x_i + b)$

Dann $\begin{cases} \alpha_i \geq 0 \\ y_i(w_i \cdot x_i + b) \geq 0 \\ \alpha_i(y_i(w_i \cdot x_i + b) = 0 \end{cases}$ (KKT)

und $\nabla_w L(w,b,\alpha) \Rightarrow w - \sum_{i=1}^{n} \alpha_i y_i x_i = 0 \Leftrightarrow w = \sum_{i=1}^{n} \alpha_i y_i x_i$ 1)

$\nabla_b L(w,b,\alpha) \Rightarrow -\sum_{i=1}^{N} \alpha_i y_i = 0 \quad \Leftrightarrow \sum_{i=1}^{n} \alpha_i y_i = 0$ 2)

in L.

$\min_{w,b} \quad L = -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$

2) $\max_{\alpha} \; -\frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i \cdot x_j)$ s.t $\sum_{i=1}^{N} \alpha_i y_i = 0$.

with $-$. $\Rightarrow$

$-(\min_{\alpha} \frac{1}{2} \sum_{i=1}^{N} \sum_{j=1}^{N} \alpha_i \alpha_j y_i y_j (x_i - x_j))$ mit KKT und 1), 2)

$y_j(w^* \cdot x_j + b^*) = 0$

$\Leftrightarrow$ muss $b = 0$, $w = 0$ $\Rightarrow$ w, b not work. $\Leftrightarrow$ Margin distance barely invisible, invalid division

---

replace 0 by 0.5

w and b work but not better than 1.