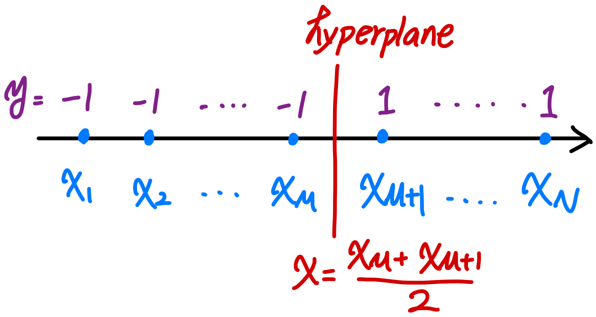
$\boxed{\text{P1.A}}$ 考慮 1D



$g_{svm} = sign(wx+b)$

顯然當 $w=1,\ b=-\frac{x_u+x_{u+1}}{2}$，所有 $x_1 \sim x_N$ 皆能正確分類

$\underline{g_{svm} = sign\left(x_i - \frac{x_u+x_{u+1}}{2}\right)}$，故選[a]

$\boxed{P2.D}$ 證明 $\|w\|^{-1}$ 是 margin distance

By L10, P10 的距離公式：$\text{distance} = \frac{1}{\|w\|} \cdot y_n(w^T z_n + b)$

By KKT condition：$b = y_n - w^T z_n$ 代入上式得：

$$\Rightarrow \text{distance} = \frac{1}{\|w\|} \cdot (w^T z_n + y_n - w^T z_n) y_n = \frac{y_n^2}{\|w\|} = \underline{\frac{1}{\|w\|}} \#$$ 故選[3]

證明 $\left(2\sum_{n=1}^{N}\alpha_n - \|\sum_{n=1}^{N}\alpha_n y_n z_n\|^2\right)^{-1/2}$

By L10, P31 的 KKT condition

$$\max_{\text{all } \alpha_n > 0,\ \underbrace{\sum y_n \alpha_n = 0}_{\text{條件①}},\ \underbrace{w = \sum \alpha_n y_n z_n}_{\text{條件②}}} -\frac{1}{2}\|\sum_{n=1}^{N}\alpha_n y_n z_n\|^2 + \sum_{n=1}^{N}\alpha_n$$

代入條件②

$$\Rightarrow -\frac{1}{2}\|w\|^2 + \sum_{n=1}^{N}\alpha_n$$

條件②

代入 $w = \sum \alpha_n y_n z_n$

$$\Rightarrow \frac{1}{2}w^T w + \sum_{n=1}^{N}\alpha_n - w^T \boxed{w}$$

$$\Rightarrow \frac{1}{2}w^T w + \sum_{n=1}^{N}\alpha_n - \sum_{n=1}^{N}\alpha_n y_n w^T z_n$$

$$\Rightarrow \frac{1}{2}w^T w + \sum_{n=1}^{N}\alpha_n \left(1 - \sum_{n=1}^{N} y_n w^T z_n\right)$$

代入條件②

$$\Rightarrow \frac{1}{2}w^T w + \sum_{n=1}^{N}\alpha_n \left(1 - \sum_{n=1}^{N} y_n (w^T z_n + b)\right)$$

代入 complementary slackness：$\alpha_n(1 - y_n(w^T z_n + b)) = 0$

由於只考慮 SV，故 $\underline{1 - y_n(w^T z_n + b) = 0}$

$$\Rightarrow \frac{1}{2}\|w\|^2$$

由前頁推導可知：

$$-\frac{1}{2}\left\|\sum_{n=1}^{N}\alpha_n y_n z_n\right\|^2+\sum_{n=1}^{N}\alpha_n=\frac{1}{2}\|w\|^2$$

$$\Rightarrow \underline{\left[2\sum_{n=1}^{N}\alpha_n-\left\|\sum_{n=1}^{N}\alpha_n y_n z_n\right\|^2\right]^{-1/2}=\|w\|^{-1}}=\text{margin distance}$$ 故选[6] #

式①

證明 $\left(\sum_{n=1}^{N}\alpha_n\right)^{-1/2}$

根據前面推導的式①

代入 $\sum\alpha_n y_n z_n=w$

$$\left[2\sum_{n=1}^{N}\alpha_n-\|\underline{\underline{\sum_{n=1}^{N}\alpha_n y_n z_n}}\|^2\right]^{-1/2}=\|w\|^{-1}$$

$$\Rightarrow 2\sum_{n=1}^{N}\alpha_n-\|w\|^2=\|w\|^2$$

$$\Rightarrow \sum_{n=1}^{N}\alpha_n=\|w\|^2$$

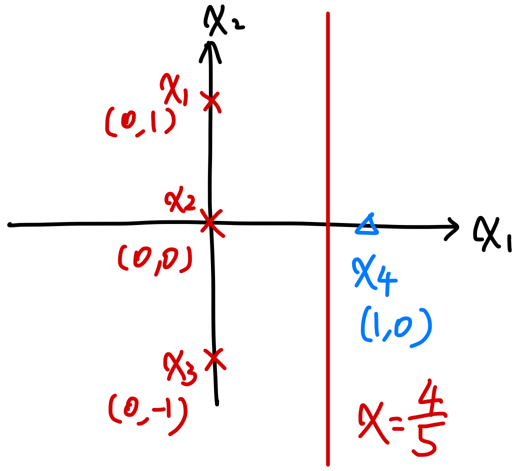
$$\Rightarrow \underline{\left(\sum_{n=1}^{N}\alpha_n\right)^{-1/2}=\|w\|^{-1}}=\text{margin distance}$$ #

$\boxed{P3\ C}$ 代入選項C，限制式變成 $x_n = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$

if $y_n > 0$, $(w^T x_n + b) \geq 1 \Rightarrow 5x_1 - 4 \geq 1 \Rightarrow \underline{x_1 \geq 1}$

if $y_n < 0$, $-(w^T x_n + b) \geq \rho_- \Rightarrow -5x_1 + 4 \geq 4 \Rightarrow \underline{x_1 \leq 0}$.

顯然滿足題目給的 $x_1 \sim x_4$



$\boxed{\text{P4.C}}$ By P3 的 primal SVM: $\min_{w,b} \frac{1}{2}w^Tw$, s.t. $\begin{cases} (w^Tx_n+b)\geq 1,\ y_n>0 \\ -(w^Tx_n+b)\geq \rho_-,\ y_n<0 \end{cases}$

By Slide L10, P.26

$$\mathcal{L}(b,w,\alpha)=\frac{1}{2}w^Tw+\sum_{n=1}^{N}\alpha_n(1-w^Tx_n-b)\cdot[\![y_n=1]\!]$$

$$+\sum_{n=1}^{N}\alpha_n(\rho_-+w^Tx_n+b)\cdot[\![y_n=-1]\!]$$

$$\frac{\mathcal{L}(b,w,\alpha)}{\partial b}=\underbrace{-\sum_{n=1}^{N}\alpha_n[\![y_n=1]\!]+\sum_{n=1}^{N}\alpha_n[\![y_n=-1]\!]=0}_{\text{condition ①}}\ \cdots 代入上式,b被消掉$$

$$\Rightarrow \min_{b,w}\frac{1}{2}w^Tw+\sum_{n=1}^{N}\alpha_n(1-w^Tx_n)[\![y_n=1]\!]+\sum_{n=1}^{N}\alpha_n(\rho_-+w^Tx_n)[\![y_n=-1]\!]$$

$$\frac{\mathcal{L}(b,w,\alpha)}{\partial w_i}=w_i-\sum_{n=1}^{N}\alpha_n x_n[\![y_n=1]\!]+\sum_{n=1}^{N}\alpha_n x_n[\![y_n=-1]\!]=0$$

$$\Rightarrow \underbrace{w=\sum_{n=1}^{N}\alpha_n x_n[\![y_n=1]\!]-\sum_{n=1}^{N}\alpha_n x_n[\![y_n=-1]\!]=0}_{\text{condition ②}}$$

代入 $\mathcal{L}(b,w,\alpha)$

$\max\left(\min_{b,w}\mathcal{L}(b,w,\alpha)\right)\leftarrow$ 代入 condition ①,②,為求过程簡化下面不特别写出限制條件

$$\Rightarrow \max\left[\frac{1}{2}w^Tw+\left(\sum_{n=1}^{N}\alpha_n[\![y_n=1]\!]-w^Tw\right)+\sum_{n=1}^{N}\alpha_n\rho_-[\![y_n=-1]\!]-w^Tw\right]$$

$$\Rightarrow \max\left[-\frac{1}{2}w^Tw+\sum_{n=1}^{N}\alpha_n[\![y_n=1]\!]+\sum_{n=1}^{N}\alpha_n\rho_-[\![y_n=-1]\!]\right]$$

$$\Rightarrow \min\left[+\frac{1}{2}w^Tw-\sum_{n=1}^{N}\alpha_n[\![y_n=1]\!]-\sum_{n=1}^{N}\alpha_n\rho_-[\![y_n=-1]\!]\right]$$

後頁續

$$\Rightarrow \min\left[+\frac{1}{2}w^T w - \sum_{n=1}^{N}\alpha_n [\![y_n = 1]\!] - \sum_{n=1}^{N}\alpha_n \rho_- [\![y_n = -1]\!]\right]$$

$$\Rightarrow \min\left[\frac{1}{2}\|\alpha_n y_n x_n\|^2 - \sum_{n=1}^{N}\alpha_n [\![y_n = 1]\!] - \sum_{n=1}^{N}\alpha_n \rho_- [\![y_n = -1]\!]\right]$$

對應題目選項[C]

$\boxed{P5.D}$ 若只考慮Support Vector，分二種情況 $y_n=1$ or $y_m=-1$

even SVM $\Rightarrow \begin{cases} w^T x_n + b = 1, \text{ if } y_n = 1 & \cdots ① \\ w^T x_m + b = -1, \text{ if } y_m = -1 & \cdots ② \end{cases} \Rightarrow ① - ② \Rightarrow \underline{w^T(x_n - x_m) = 2} \cdots ⑤$

uneven SVM $\Rightarrow \begin{cases} w'^T x_n + b' = 1, \text{ if } y_n = 1 & \cdots ③ \\ w'^T x_m + b' = -\rho_-, \text{ if } y_m = -1 & \cdots ④ \end{cases} \Rightarrow ③ - ④ \Rightarrow \underline{w'^T(x_n - x_m) = 1 + \rho_-} \cdots ⑥$

$$\frac{⑤}{⑥} \Rightarrow \frac{w^T(x_n - x_m)}{w'^T(x_n - x_m)} = \frac{2}{1+\rho_-} \Rightarrow \underline{w' = \frac{1+\rho_-}{2} w}$$

By KKT condition: $w = \sum_{n=1}^{N} \alpha_n y_n x_n$ 代入上式

$$\Rightarrow \sum_{n=1}^{N} \alpha_n' \cancel{y_n x_n} = \frac{1+\rho_-}{2} \left(\sum_{n=1}^{N} \alpha_n \cancel{y_n x_n} \right)$$

$$\Rightarrow \sum_{n=1}^{N} \alpha_n' = \underline{\frac{1+\rho_1}{2}} \sum_{n=1}^{N} \alpha_n$$

故選[D]

$\boxed{P6, A}$ $K(x, x') = (x^T x')^Q$，代表$\Phi(x)$只留最高次項

$\Phi(x) = (x_1^Q, x_2^Q, \cdots, x_d^Q, x_1 x_2^{Q-1}, x_1 x_3^{Q-1}, \cdots\cdots x_1 x_d^{Q-1}, \cdots\cdots$

我們可以想像有Q個相同的球，要放進d個箱子裡，一共有幾種不同的放法？

$\Rightarrow$ |O|O||O ⋯ O||

一共有Q個球，d-1個bar做排列 $\Rightarrow \dfrac{(Q+d-1)!}{Q!(d-1)!} \Rightarrow \underline{\dbinom{Q+d-1}{Q}}$ # 故選[A]

此為不全相異排列

$\boxed{P7.D}$ By Slide L11, P7: $K_2(x, x') = 1 + 2x^Tx' + (x^Tx')^2$

$= (1 + x^Tx')^2$

where $\Phi_2(x) = (1, \sqrt{2}x_1, \sqrt{2}x_2, \cdots, \sqrt{2}x_d, x_1^2, x_2^2, \cdots, x_d^2)$

令 $x = (1, 0, 0, \cdots, 0)$, $x' = (-1, 0, 0, \cdots, 0)$, 此情形會出現 upper bound.

$\Phi(x) - \Phi(x') = (\cancel{1}, \sqrt{2}x_1, \cancel{\sqrt{2}x_2}, \cancel{\cdots}, \cancel{\sqrt{2}x_d}, \cancel{x_1^2}, \cancel{x_2^2}, \cdots, \cancel{x_d^2})$

$- (\cancel{1}, \sqrt{2}x_1', \cancel{\sqrt{2}x_2'}, \cancel{\cdots}, \cancel{\sqrt{2}x_d'}, \cancel{x_1'^2}, \cancel{x_2'^2}, \cdots, \cancel{x_d'^2})$

$\Rightarrow \sqrt{2} - (-\sqrt{2}) = 2\sqrt{2}$

$\Rightarrow \|\Phi(x) - \Phi(x')\|^2 = (2\sqrt{2})^2 = \underline{8}$ #

$\boxed{\text{P8.C}}$ PLA更新公式：$w_{t+1} \leftarrow w_t + y_{n(t)} \phi(x_{n(t)})$

所以 $\alpha_{t+1}[n(t)] = y_{n(t)}$，才能達到更新的效果

P9. A

Primal SVM 根據題目:

$$\min_{w,b,\xi} \frac{1}{2}w^Tw + \sum_{n=1}^{N} u_n \cdot \xi_n \quad , \text{s.t. } y_n(w^T\Phi(x_n)+b) \geq 1-\xi_n$$
$$\xi \geq 0, \ n = 1, \dots, N$$

By Slide L11. P27 推導成 dual

$$\mathcal{L}(b,w,\xi,\alpha,\beta,u_n) = \frac{1}{2}w^Tw + \sum_{n=1}^{N} u_n \cdot \xi_n + \sum_{n=1}^{N} \alpha_n \cdot (1-\xi_n - y_n(w^Tz_n+b)) + \sum_{n=1}^{N} \beta_n \cdot (-\xi_n)$$

$$\frac{\partial \mathcal{L}}{\partial \xi_n} = 0 \Rightarrow \underline{u_n - \alpha_n - \beta_n = 0} \quad \text{condition}$$

$$\max\left(\min_{b,w,\xi} \mathcal{L}(b,w,\xi,\alpha,\beta,u_n)\right)$$

$$\Rightarrow \max\left(\min_{b,w,\xi} \frac{1}{2}w^Tw + \sum_{n=1}^{N} \alpha_n(1-y_n(w^Tz_n+b))\right) \cancel{+\sum_{n=1}^{N}(u_n-\alpha_n-\beta_n)\cdot\xi_n}$$

$\beta_n = u_n - \alpha_n$

u_n 在這個步驟被消掉,成為 condition,代表後面的推導也跟講義的 soft-margin SVM dual problem 一樣

By Slide L11, P30,

$$\underline{\min_{\alpha} \frac{1}{2}\sum\sum \alpha_n\alpha_m y_n y_m z_n^T z_m - \sum_{n=1}^{N}\alpha_n} \ \#$$ 故選 [a]

$\boxed{\text{P10, B}}$

$$\int_0^1 \left(\frac{1}{2}(1-\rho)^2-(1-\rho)\right)^2\cdot\rho\, d\rho$$

$$=\int_0^1\left(\frac{1}{2}(\rho^2-2\rho+1)-1+\rho\right)^2\cdot\rho\,d\rho$$

$$=\int_0^1\left(\frac{\rho^2-1}{2}\right)^2\cdot\rho\,d\rho$$

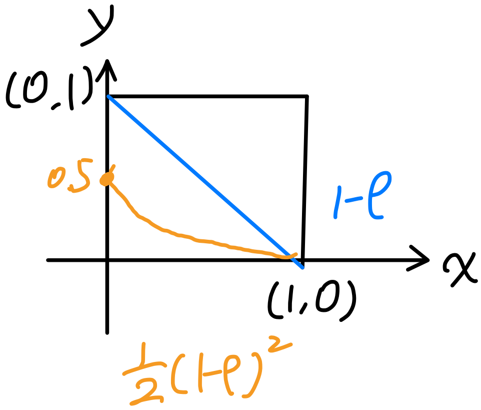
$$=\int_0^1\left(\frac{\rho^4-2\rho^2+1}{4}\right)\cdot\rho$$

$$=\int_0^1\frac{\rho^5-2\rho^3+\rho}{4}$$

$$=\frac{1}{4}\left[\frac{\rho^6}{6}-\frac{2\rho^4}{4}+\frac{\rho^2}{2}\right]\Bigg|_0^1$$

$$=\frac{1}{4}\left(\frac{1}{6}-\frac{1}{2}+\frac{1}{2}\right)$$

$=\underline{\frac{1}{24}}$ # 故选[B]



PILA

```
./svm-train -s 0 -t 0 -c 10 satimage_5vo.scale
```

```python
# Transform data into '5' versus 'not 5'
#
s = ""
with open('libsvm/satimage.scale.t', 'r') as f:
    for line in f.readlines():
        if line[0] != '6':
            s += '1' + line[1:] # "not X"
        else:
            s += '2' + line[1:] # "is X"

with open('libsvm/satimage_6vo.scale.t', 'w') as f:
    f.write(s)
```

```python
from collections import defaultdict
import math

ans = defaultdict(int)
sum_alpha = 0

with open("satimage_5vo.scale.model", 'r') as f:
    for line in f.readlines():
        if line.find(':') != -1: # It's a SV
            l_list = line.split("\n")[0].rstrip().split(" ")

            y_alpha = float(l_list[0])
            sum_alpha += abs(y_alpha)
            for element in l_list[1:]:
                idx, value = element.split(":")
                ans[int(idx)] += float(value)*y_alpha

# Get |w|
l_acc = 0
for i in ans:
    l_acc += ans[i]**2
w_len = math.sqrt(l_acc)
print(f"|w| = {w_len}")
```

P12.C

```
echo "============ 2 vs. not 2 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_2vo.scale
./svm-predict satimage_2vo.scale satimage_2vo.scale.model output.txt
echo "============ 3 vs. not 3 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_3vo.scale
./svm-predict satimage_3vo.scale satimage_3vo.scale.model output.txt
echo "============ 4 vs. not 4 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_4vo.scale
./svm-predict satimage_4vo.scale satimage_4vo.scale.model output.txt
echo "============ 5 vs. not 5 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_5vo.scale
./svm-predict satimage_5vo.scale satimage_5vo.scale.model output.txt
echo "============ 6 vs. not 6 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_6vo.scale
./svm-predict satimage_6vo.scale satimage_6vo.scale.model output.txt
```

PB.E

```
echo "============ 2 vs. not 2 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_2vo.scale
./svm-predict satimage_2vo.scale satimage_2vo.scale.model output.txt
echo "============ 3 vs. not 3 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_3vo.scale
./svm-predict satimage_3vo.scale satimage_3vo.scale.model output.txt
echo "============ 4 vs. not 4 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_4vo.scale
./svm-predict satimage_4vo.scale satimage_4vo.scale.model output.txt
echo "============ 5 vs. not 5 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_5vo.scale
./svm-predict satimage_5vo.scale satimage_5vo.scale.model output.txt
echo "============ 6 vs. not 6 =============="
./svm-train -s 0 -t 1 -d 3 -c 10 -r 1 -g 1 satimage_6vo.scale
./svm-predict satimage_6vo.scale satimage_6vo.scale.model output.txt
```

P14, D

```
echo "============ C = 0.01 ==============="
./svm-train -s 0 -t 2 -g 10 -c 0.01 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ C = 0.1 ==============="
./svm-train -s 0 -t 2 -g 10 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ C = 1 ==============="
./svm-train -s 0 -t 2 -g 10 -c 1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ C = 10 ==============="
./svm-train -s 0 -t 2 -g 10 -c 10 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ C = 100 ==============="
./svm-train -s 0 -t 2 -g 10 -c 100 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
```

P15.B

```
echo "============ r = 0.1 =============="
./svm-train -s 0 -t 2 -g 0.1 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ r = 1 =============="
./svm-train -s 0 -t 2 -g 1 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ r = 10 =============="
./svm-train -s 0 -t 2 -g 10 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ r = 100 =============="
./svm-train -s 0 -t 2 -g 100 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
echo "============ r = 1000 =============="
./svm-train -s 0 -t 2 -g 1000 -c 0.1 satimage_1vo.scale
./svm-predict satimage_1vo.scale.t satimage_1vo.scale.model output.txt
```

P16, A

```
# echo "============ r = 0.1 ==============="
# ./svm-train -s 0 -t 2 -g 0.1 -c 0.1 -v 22.175 satimage_1vo.scale
# echo "============ r = 1 ==============="
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
./svm-train -s 0 -t 2 -g 1 -c 0.1 -v 22.175 satimage_1vo.scale
# echo "============ r = 10 ==============="
# ./svm-train -s 0 -t 2 -g 10 -c 0.1 -v 22.175 satimage_1vo.scale
# echo "============ r = 100 ==============="
# ./svm-train -s 0 -t 2 -g 100 -c 0.1 -v 22.175 satimage_1vo.scale
# echo "============ r = 1000 ==============="
# ./svm-train -s 0 -t 2 -g 1000 -c 0.1 -v 22.175 satimage_1vo.scale
```