Расчётка по SVM для датасета Wine (2 признака)

1. Исходные данные

Точки $x_i = (\text{param\_1}, \text{param\_2})$, метки $y_i \in \{-1, +1\}$:

param_1	param_2	y
0.27	0.36	1
0.30	0.34	1
0.70	0.00	-1
0.88	0.00	-1

Расчёт средних μ и стандартных отклонений σ

Дано: четыре точки $x_i = (x_{i1}, x_{i2})$:

$$x_1 = (0.27,\, 0.36), \quad x_2 = (0.30,\, 0.34), \quad x_3 = (0.70,\, 0.00), \quad x_4 = (0.88,\, 0.00).$$

Средние (по формуле $\mu_j = \frac{1}{n}\sum_i x_{ij}$, $n = 4$):

$$\mu_1 = \frac{0.27 + 0.30 + 0.70 + 0.88}{4} = \frac{2.15}{4} = 0.53750000,$$
$$\mu_2 = \frac{0.36 + 0.34 + 0.00 + 0.00}{4} = \frac{0.70}{4} = 0.17500000.$$

Дисперсии (population): $\sigma_j^2 = \frac{1}{n}\sum_i (x_{ij} - \mu_j)^2$.

$$\begin{aligned}
\sigma_1^2 &= \frac{(0.27 - 0.5375)^2 + (0.30 - 0.5375)^2 + (0.70 - 0.5375)^2 + (0.88 - 0.5375)^2}{4} \\
&= \frac{(-0.2675)^2 + (-0.2375)^2 + (0.1625)^2 + (0.3425)^2}{4} \\
&= \frac{0.07155625 + 0.05640625 + 0.02640625 + 0.11730625}{4} \\
&= \frac{0.27167500}{4} = 0.06791875,
\end{aligned}$$

$$\begin{aligned}
\sigma_2^2 &= \frac{(0.36 - 0.175)^2 + (0.34 - 0.175)^2 + (0.00 - 0.175)^2 + (0.00 - 0.175)^2}{4} \\
&= \frac{(0.185)^2 + (0.165)^2 + (-0.175)^2 + (-0.175)^2}{4} \\
&= \frac{0.03422500 + 0.02722500 + 0.03062500 + 0.03062500}{4} \\
&= \frac{0.12270000}{4} = 0.03067500.
\end{aligned}$$

Стандартные отклонения (квадратные корни из дисперсий):

$$\sigma_1 = \sqrt{0.06791875} = 0.26061226, \qquad \sigma_2 = \sqrt{0.03067500} = 0.17514280.$$

Итого:

$$\boxed{\mu = (0.53750000,\ 0.17500000), \qquad \sigma = (0.26061226,\ 0.17514280)}.$$

Нормализация признаков (до/после и поэлементные расчёты)

Стандартизируем каждый признак по формуле (population std, как в C++):

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \qquad \sigma_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\mu_j)^2}, \qquad z_{ij} = \frac{x_{ij}-\mu_j}{\max(\sigma_j, 10^{-12})}.$$

Для текущего набора ($n = 4$):

$$\mu = (0.53750000,\ 0.17500000), \qquad \sigma = (0.26061226,\ 0.17514280).$$

Таблица "до/после": исходные x_{ij} и стандартизованные z_{ij}

i	x_{i1}	x_{i2}	z_{i1}	z_{i2}
1	0.27	0.36	−1.02642907	1.05628094
2	0.30	0.34	−0.91131553	0.94208840
3	0.70	0.00	0.62353168	−0.99918467
4	0.88	0.00	1.31421292	−0.99918467

Поэлементные расчёты z_{ij} (с подстановкой)

$$\begin{aligned}
z_{11} &= \frac{x_{11}-\mu_1}{\sigma_1} \\
&= \frac{0.27-0.5375}{0.26061226} \\
&= \frac{-0.2675}{0.26061226} = -1.02642907, \\
z_{12} &= \frac{x_{12}-\mu_2}{\sigma_2} \\
&= \frac{0.36-0.175}{0.17514280} \\
&= \frac{0.185}{0.17514280} = 1.05628094, \\
z_{21} &= \frac{x_{21}-\mu_1}{\sigma_1} \\
&= \frac{0.30-0.5375}{0.26061226} \\
&= \frac{-0.2375}{0.26061226} = -0.91131553, \\
z_{22} &= \frac{x_{22}-\mu_2}{\sigma_2} \\
&= \frac{0.34-0.175}{0.17514280} \\
&= \frac{0.165}{0.17514280} = 0.94208840.
\end{aligned}$$

$$\begin{aligned}
z_{31} &= \frac{x_{31}-\mu_1}{\sigma_1} \\
&= \frac{0.70-0.5375}{0.26061226} \\
&= \frac{0.1625}{0.26061226} = 0.62353168, \\
z_{32} &= \frac{x_{32}-\mu_2}{\sigma_2} \\
&= \frac{0.00-0.175}{0.17514280} \\
&= \frac{-0.175}{0.17514280} = -0.99918467, \\
z_{41} &= \frac{x_{41}-\mu_1}{\sigma_1} \\
&= \frac{0.88-0.5375}{0.26061226} \\
&= \frac{0.3425}{0.26061226} = 1.31421292, \\
z_{42} &= \frac{x_{42}-\mu_2}{\sigma_2} \\
&= \frac{0.00-0.175}{0.17514280} \\
&= \frac{-0.175}{0.17514280} = -0.99918467.
\end{aligned}$$

2. Нормализация (standardize как в C++)

Для каждого признака j берём среднее и стандартное отклонение (population std):

$$\mu_j = \frac{1}{n}\sum_{i=1}^{n} x_{ij}, \qquad \sigma_j = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_{ij}-\mu_j)^2}.$$

Стандартизованные признаки: $z_{ij} = \dfrac{x_{ij}-\mu_j}{\max(\sigma_j, 10^{-12})}$. Численно для нашего набора ($n=4$):

$$\mu = (0.537500,\ 0.175000), \qquad \sigma = (0.260612,\ 0.175143).$$

3. Стандартизованные векторы Z_i

$$\begin{aligned}
Z_1 &= (-1.02642907,\ 1.05628094), \\
Z_2 &= (-0.91131553,\ 0.94208840), \\
Z_3 &= (0.62353168,\ -0.99918467), \\
Z_4 &= (1.31421292,\ -0.99918467).
\end{aligned}$$

4. Матрица Грама $K = ZZ^\top$

Определение: $K_{ij} = \langle Z_i, Z_j \rangle = \sum_{k=1}^{2} Z_{ik}\, Z_{jk}$.

$$K = \begin{bmatrix} 2.16928606 & 1.93051078 & -1.69543077 & -2.40436607 \\ 1.93051078 & 1.71802656 & -1.50955440 & -2.13898294 \\ -1.69543077 & -1.50955440 & 1.38716176 & 1.81782340 \\ -2.40436607 & -2.13898294 & 1.81782340 & 2.72552562 \end{bmatrix}.$$

Матрица Грама K для стандартизованных векторов

По определению:

$$K_{ij} = \langle Z_i, Z_j \rangle = \sum_{k=1}^{2} Z_{ik}\, Z_{jk}.$$

Векторы Z_i (из нормализации):

$$\begin{aligned} Z_1 &= (-1.02642907,\ \ 1.05628094), \\ Z_2 &= (-0.91131553,\ \ 0.94208840), \\ Z_3 &= (\ \ 0.62353168,\ -0.99918467), \\ Z_4 &= (\ \ 1.31421292,\ -0.99918467). \end{aligned}$$

Пошаговые вычисления всех элементов K_{ij}

По формуле $K_{ij} = \langle Z_i, Z_j \rangle = \sum_{k=1}^{2} Z_{ik} Z_{jk}$ при

$$\begin{aligned} Z_1 &= (-1.02642907,\ \ 1.05628094), \\ Z_2 &= (-0.91131553,\ \ 0.94208840), \\ Z_3 &= (\ \ 0.62353168,\ -0.99918467), \\ Z_4 &= (\ \ 1.31421292,\ -0.99918467), \end{aligned}$$

получаем:

Строка 1: $K_{11} = (-1.02642907)(-1.02642907) + (1.05628094)(1.05628094) = 2.16928606,$
$K_{12} = (-1.02642907)(-0.91131553) + (1.05628094)(0.94208840) = 1.93051078,$
$K_{13} = (-1.02642907)(0.62353168) + (1.05628094)(-0.99918467) = -1.69543077,$
$K_{14} = (-1.02642907)(1.31421292) + (1.05628094)(-0.99918467) = -2.40436607;$

Строка 2: $K_{21} = (-0.91131553)(-1.02642907) + (0.94208840)(1.05628094) = 1.93051078,$
$K_{22} = (-0.91131553)(-0.91131553) + (0.94208840)(0.94208840) = 1.71802656,$
$K_{23} = (-0.91131553)(0.62353168) + (0.94208840)(-0.99918467) = -1.50955440,$
$K_{24} = (-0.91131553)(1.31421292) + (0.94208840)(-0.99918467) = -2.13898294;$

Строка 3: $K_{31} = (0.62353168)(-1.02642907) + (-0.99918467)(1.05628094) = -1.69543077,$
$K_{32} = (0.62353168)(-0.91131553) + (-0.99918467)(0.94208840) = -1.50955440,$
$K_{33} = (0.62353168)(0.62353168) + (-0.99918467)(-0.99918467) = 1.38716176,$
$K_{34} = (0.62353168)(1.31421292) + (-0.99918467)(-0.99918467) = 1.81782340;$

Строка 4: $K_{41} = (1.31421292)(-1.02642907) + (-0.99918467)(1.05628094) = -2.40436607,$
$K_{42} = (1.31421292)(-0.91131553) + (-0.99918467)(0.94208840) = -2.13898294,$
$K_{43} = (1.31421292)(0.62353168) + (-0.99918467)(-0.99918467) = 1.81782340,$
$K_{44} = (1.31421292)(1.31421292) + (-0.99918467)(-0.99918467) = 2.72552562.$

$$K = \begin{bmatrix} 2.16928606 & 1.93051078 & -1.69543077 & -2.40436607 \\ 1.93051078 & 1.71802656 & -1.50955440 & -2.13898294 \\ -1.69543077 & -1.50955440 & 1.38716176 & 1.81782340 \\ -2.40436607 & -2.13898294 & 1.81782340 & 2.72552562 \end{bmatrix}.$$

Мини-проверка: $\|Z_2 - Z_3\|^2$ через элементы K и вычисление a

По определению,

$$\|Z_2 - Z_3\|^2 = \langle Z_2 - Z_3,\ Z_2 - Z_3 \rangle = K_{22} - 2K_{23} + K_{33}.$$

Подставим числовые значения:

$$\begin{aligned} \|Z_2 - Z_3\|^2 &= 1.71802656 - 2(-1.50955440) + 1.38716176 \\ &= 1.71802656 + 3.01910880 + 1.38716176 \\ &= \boxed{6.12429712}. \end{aligned}$$

По ККТ для двух опорных векторов Z_2, Z_3:

$$a\,\|Z_2 - Z_3\|^2 = 2 \quad \Longrightarrow \quad a = \frac{2}{\|Z_2 - Z_3\|^2} = \frac{2}{6.12429712} = \boxed{0.326606}\ (\approx 0.3266).$$

Прямая задача (жёсткий зазор): смысл и цель

Пусть разделяющая гиперплоскость задаётся $w^\top x + b = 0$. Знак $f(x) = \operatorname{sign}(w^\top x + b)$ — предсказанный класс.

Функциональный маржин точки (x_i, y_i):

$$m_i^{\text{func}} = y_i\,(w^\top x_i + b).$$

Геометрический маржин — это расстояние до гиперплоскости с правильным знаком:

$$m_i^{\text{geom}} = \frac{y_i\,(w^\top x_i + b)}{\|w\|}.$$

Максимизировать минимальный геометрический маржин эквивалентно задаче: зафиксировать масштаб так, чтобы $\min_i m_i^{\text{func}} = 1$ (то есть точки «на границе» имеют $y_i(w^\top x_i + b) = 1$), и затем *минимизировать* норму w:

$$\boxed{\min_{w,b}\ \tfrac{1}{2}\|w\|^2 \quad \text{при} \quad y_i\,(w^\top x_i + b)\ \geq\ 1,\ \forall i.}$$

Почему это про «максимизацию зазора»? Ширина зазора (двойная полоса между опорными гиперпрямыми) равна

$$\text{width} = \frac{2}{\|w\|}.$$

Значит, минимизация $\|w\|$ *при сохранении правильной классификации с запасом 1* даёт максимальную ширину зазора.

Опорные векторы — это точки, для которых ограничение активно (равенство $y_i(w^\top x_i + b) = 1$). Они «держат» оптимальную гиперплоскость; точки дальше зазора (> 1) на решение не влияют.

KKT-условия для оптимума:

$$\begin{aligned}
&y_i(w^\top x_i + b) \geq 1, \quad \alpha_i \geq 0,\\
&\alpha_i\big(1 - y_i(w^\top x_i + b)\big) = 0,\\
&w = \sum_i \alpha_i y_i x_i, \quad \sum_i \alpha_i y_i = 0.
\end{aligned}$$

Они формализуют: у неопорных $\alpha_i = 0$ и $m_i^{\text{func}} > 1$; у опорных $\alpha_i > 0$ и $m_i^{\text{func}} = 1$.

SVM с мягким зазором (soft–margin): зачем и как работает

Зачем. Если классы не линейно разделимы или данные зашумлены, требование $y_i(w^\top x_i + b) \geq 1$ для всех i невыполнимо. Разрешим нарушения, но будем их штрафовать. Это делает модель устойчивой и позволяющей компромисс «ширина зазора ошибки».

Прямая (primal) постановка с допусками

Вводим неотрицательные допуски (слэки) $\xi_i \geq 0$ и параметр $C > 0$:

$$\boxed{\min_{w,b,\xi}\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\xi_i \quad \text{при} \quad y_i(w^\top x_i + b) \geq 1 - \xi_i, \quad \xi_i \geq 0\ \forall i.}$$

Эквивалентно (через hinge-loss):

$$\min_{w,b}\ \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{n}\max\bigl(0,\ 1 - y_i(w^\top x_i + b)\bigr).$$

Интерпретация C: чем больше C, тем сильнее штраф за нарушения (меньше допусков, уже зазор); чем меньше C, тем шире зазор и больше допустимых нарушений.

Двойственная (dual) постановка

Строим лагранжиан с множителями $\alpha_i \geq 0$ для ограничений марджина и $\rho_i \geq 0$ для $\xi_i \geq 0$:

$$\mathcal{L} = \tfrac{1}{2}\|w\|^2 + C\sum_i \xi_i - \sum_i \alpha_i\bigl(y_i(w^\top x_i + b) - 1 + \xi_i\bigr) - \sum_i \rho_i \xi_i.$$

Условия стационарности:

$$\frac{\partial \mathcal{L}}{\partial w} = 0 \Rightarrow w = \sum_i \alpha_i y_i x_i, \qquad \frac{\partial \mathcal{L}}{\partial b} = 0 \Rightarrow \sum_i \alpha_i y_i = 0,$$

$$\frac{\partial \mathcal{L}}{\partial \xi_i} = 0 \Rightarrow C - \alpha_i - \rho_i = 0\ \Rightarrow\ 0 \leq \alpha_i \leq C.$$

Подставляя обратно, получаем dual-задачу (как у hard–margin, но с *боковыми* ограничениями):

$$\boxed{\max_{\alpha}\ \sum_{i=1}^{n}\alpha_i - \tfrac{1}{2}\sum_{i,j=1}^{n}\alpha_i\alpha_j y_i y_j K(x_i, x_j) \quad \text{при} \quad \sum_i \alpha_i y_i = 0,\ \ 0 \leq \alpha_i \leq C.}$$

(Для линейного SVM $K(x_i, x_j) = \langle x_i, x_j\rangle$; с ядром — любой допустимый K.)

KKT-условия и геометрическая интерпретация

Помимо стационарности действуют дополняющие нежёсткости:

$$\alpha_i\bigl(1 - y_i(w^\top x_i + b) - \xi_i\bigr) = 0, \qquad \rho_i\,\xi_i = 0, \qquad \xi_i \geq 0,\ \alpha_i \geq 0,\ \rho_i \geq 0.$$

Введём сокращение $f_i := y_i(w^\top x_i + b)$. Тогда практические правила удобно записать так:

$$\begin{aligned} \alpha_i = 0 \ &\Rightarrow\ f_i > 1 && \text{(вне зазора)},\\ 0 < \alpha_i < C \ &\Rightarrow\ f_i = 1,\ \xi_i = 0 && \text{(на границе)},\\ \alpha_i = C \ &\Rightarrow\ f_i \leq 1 && \text{(внутри зазора/ошибка)}. \end{aligned}$$

Дополнительно: $\xi_i = \max(0,\,1 - f_i)$, поэтому

$$f_i < 1\ \Rightarrow\ \xi_i > 0, \qquad f_i < 0\ \Rightarrow\ \xi_i > 1 \text{ (ошибка классификации)}.$$

Восстановление w и b на практике

$$w=\sum_i \alpha_i y_i x_i, \qquad \text{а } b \text{ удобно усреднить по индексам } \mathcal{M}=\{i:\ 0<\alpha_i<C\}:$$

$$b=\frac{1}{|\mathcal{M}|}\sum_{i\in\mathcal{M}}\Big(y_i-\sum_j \alpha_j y_j K(x_j,x_i)\Big).$$

Если множество $\mathcal{M}$ пусто (бывает при сильном перекрытии и всех α_i на границе C), используют точки с $\alpha_i=C$ (внутри зазора) и соответствующие формулы/эвристики библиотек.

Замечания про масштабирование

Так как C измеряет «цену» допусков в тех же единицах, что и $\|w\|^2$, **масштаб признаков** влияет на эффективный диапазон C. Поэтому перед обучением мы стандартизируем признаки $z=(x-\mu)/\sigma$; dual остаётся тем же, но K строится по Z, а итоговые w,b при необходимости переводятся обратно в пространство x:

$$w_x=\frac{w_z}{\sigma}, \qquad b_x=b_z-w_z^\top\frac{\mu}{\sigma}.$$

Почему решение можно искать по двум опорным векторам

Интуиция. В hard–margin SVM оптимальная полоса максимальной ширины касается классов *ровно на границе* зазора. Эти точки и есть *опорные векторы* (SV). Если на каждой стороне зазора по одной точке (в общем положении), то они полностью задают ориентацию и положение гиперплоскости. Остальные точки лежат дальше зазора и на решение не влияют.

KKT–скрининг. В оптимуме:

$$\alpha_i>0 \;\Leftrightarrow\; y_i(w^\top x_i+b)=1 \quad (\text{точно на границе}), \qquad \alpha_i=0 \;\Leftrightarrow\; y_i(w^\top x_i+b)>1.$$

На наших данных в Z-пространстве *ровно* две точки имеют маржин $m_i=1$ (обозначим их Z_2 и Z_3), а у остальных $m_i>1$. Следовательно,

$$\alpha_1=\alpha_4=0, \qquad \alpha_2>0,\ \alpha_3>0.$$

Сведение dual к одной переменной. В dual-задаче

$$\max_{\alpha\ge 0}\ \sum_i \alpha_i-\tfrac{1}{2}\sum_{i,j}\alpha_i\alpha_j y_i y_j K_{ij}, \quad \text{при } \sum_i \alpha_i y_i=0,$$

положим (по KKT) $\alpha_1=\alpha_4=0$ и обозначим $\alpha_2=\alpha_3=a$. Линейное ограничение выполняется автоматически: $a\cdot(+1)+a\cdot(-1)=0$. Тогда целевая упрощается до

$$W(a)=2a-\frac{1}{2}\,a^2\big(K_{22}-2K_{23}+K_{33}\big).$$

Условие максимума: $W'(a)=0 \;\Rightarrow\; a=\dfrac{2}{K_{22}-2K_{23}+K_{33}}$.

Восстановление параметров.

$$w=\sum_i \alpha_i y_i Z_i=a(Z_2-Z_3), \qquad b=1-w^\top Z_2=1-a\,(K_{22}-K_{23}).$$

Численно для нашего K: $a\approx 0.326606$, $w_z\approx(-0.501232,\ 0.633958)$, $b_z\approx -0.054025$.

Когда двух SV *недостаточно*? Если на границе зазора лежит не одна, а *несколько* точек (коллинеарные/копланарные случаи, случаи вырожденности), то ненулевых α_i будет больше двух. Тогда либо решают полный dual (QP), либо берут активный набор SV и решают систему условий $y_i(w^\top Z_i + b) = 1$ для всех активных индексов. В общем случае минимальный набор SV, необходимый, чтобы однозначно задать гиперплоскость в $\mathbb{R}^d$, может быть больше двух (до $d+1$), но на нашем 2D-наборе в общем положении достаточно двух — по одному на каждую сторону зазора.

Решение по двум опорным векторам

Работаем в стандартизованных координатах $Z = (x-\mu)/\sigma$. Метки $y = (+1, +1, -1, -1)$, опорные векторы по проверке ККТ: Z_2 и Z_3 (у них $m_i = 1$).

1) Сведение dual-задачи к одной переменной

Двойственная задача (для линейного ядра $K_{ij} = \langle Z_i, Z_j\rangle$):

$$\max_{\alpha}\ \mathcal{W}(\alpha) = \sum_{i=1}^{4}\alpha_i - \frac{1}{2}\sum_{i,j=1}^{4}\alpha_i\alpha_j y_i y_j K_{ij}, \quad \text{при } \sum_{i=1}^{4}\alpha_i y_i = 0,\ \alpha_i \ge 0.$$

По ККТ у неопорных $\alpha_1 = \alpha_4 = 0$, у опорных положим $\alpha_2 = \alpha_3 = a$. Линейное ограничение выполнено: $a\cdot(+1) + a\cdot(-1) = 0$.

Тогда цель упрощается:

$$\mathcal{W}(a) = 2a - \frac{1}{2}\,a^2\big(K_{22} - 2K_{23} + K_{33}\big).$$

Условие максимума: $\mathcal{W}'(a) = 0 \Rightarrow \boxed{a = \dfrac{2}{K_{22} - 2K_{23} + K_{33}}}$.

2) Подстановка чисел и значение a

Из матрицы Грама (на Z):

$$K_{22} = 1.71802656, \quad K_{23} = -1.50955440, \quad K_{33} = 1.38716176.$$

Тогда

$$K_{22} - 2K_{23} + K_{33} = 1.71802656 - 2(-1.50955440) + 1.38716176 = 6.12429712,$$

$$\boxed{a = \frac{2}{6.12429712} = 0.326606}.$$

3) Восстановление w

По стационарности $w = \sum_i \alpha_i y_i Z_i = a(Z_2 - Z_3)$. Сами векторы:

$$Z_2 = (-0.91131553,\ 0.94208840), \quad Z_3 = (\ \ 0.62353168,\ -0.99918467),$$

$$Z_2 - Z_3 = (-1.53484721,\ 1.94127307).$$

Значит

$$\boxed{w_z = a(Z_2 - Z_3) = 0.326606\cdot(-1.53484721,\ 1.94127307) = (-0.501232,\ 0.633958)}.$$

4) Восстановление b (через равенство на границе)

Для опорного Z_2 верно $w^\top Z_2 + b = 1$, откуда

$$b_z = 1 - w^\top Z_2 = 1 - a\,(K_{22} - K_{23}).$$

Численно: $K_{22} - K_{23} = 1.71802656 - (-1.50955440) = 3.22758096$,

$$\boxed{b_z = 1 - 0.326606 \cdot 3.22758096 = -0.054025}.$$

(Проверка вторым СВ: $w^\top Z_3 + b = -1$ выполняется.)

5) Проверка условий ККТ и маржинов

Решающая функция в Z выражается через K:

$$f(Z_i) = w^\top Z_i + b = a\,(K_{i2} - K_{i3}) + b.$$

Считаем значения:

$$\begin{aligned}
f(Z_1) &= a(K_{12} - K_{13}) + b = 0.326606 \cdot (1.93051078 - (-1.69543077)) - 0.054025 \\
&= 0.326606 \cdot 3.62594155 - 0.054025 \approx 1.130976, \\
f(Z_2) &= a(K_{22} - K_{23}) + b = 0.326606 \cdot 3.22758096 - 0.054025 \approx 1.000000, \\
f(Z_3) &= a(K_{32} - K_{33}) + b = 0.326606 \cdot (-1.50955440 - 1.38716176) - 0.054025 \approx -1.000000, \\
f(Z_4) &= a(K_{42} - K_{43}) + b = 0.326606 \cdot (-2.13898294 - 1.81782340) - 0.054025 \approx -1.346298.
\end{aligned}$$

Функциональные маржины $m_i = y_i f(Z_i)$:

$$m_1 \approx 1.131 > 1, \quad m_2 = 1, \quad m_3 = 1, \quad m_4 \approx 1.346 > 1.$$

Итого: $\alpha_{2,3} > 0$ и $m_{2,3} = 1$ (опорные), $\alpha_{1,4} = 0$ и $m_{1,4} > 1$ — ККТ выполнены.

6) Норма и ширина зазора

$$\|w_z\| = \sqrt{(-0.501232)^2 + (0.633958)^2} \approx 0.808, \qquad \boxed{\text{width} = \frac{2}{\|w\|} \approx \frac{2}{0.808} \approx 2.475}.$$

Эквивалентно, $\text{width} = \|Z_2 - Z_3\| = \sqrt{K_{22} - 2K_{23} + K_{33}} = \sqrt{6.12429712} \approx 2.475$.

7) (Опционально) Возврат в исходное пространство x

Если нужна гиперплоскость в исходных единицах:

$$w_x = \frac{w_z}{\sigma} = (-1.923287,\ 3.619663), \qquad b_x = b_z - w_z^\top \frac{\mu}{\sigma} = 0.346301,$$

и разделяющая прямая задаётся $w_x^\top x + b_x = 0$.

Решающая функция и маржины (в стандартизованном пространстве Z)

Решающая функция: $f(Z_i) = w_z^\top Z_i + b_z = a\,(K_{i2} - K_{i3}) + b_z$. Функциональные маржины: $m_i = y_i\,f(Z_i)$.

i	$Z_i = (z_{i1}, z_{i2})$	$f(Z_i)$	y_i	$m_i = y_i f(Z_i)$	SV?
1	$(-1.026429,\ 1.056281)$	1.130976	1	1.130976	нет
2	$(-0.911316,\ 0.942088)$	1.000000	1	1.000000	да
3	$(\ \ 0.623532,\ -0.999185)$	-1.000000	-1	1.000000	да
4	$(\ \ 1.314213,\ -0.999185)$	-1.346298	-1	1.346298	нет

Параметры (в Z): $w_z = (-0.501232,\ 0.633958)$, $b_z = -0.054025$, $a = 0.326606$.

Ширина зазора

Норма весов и ширина:

$$\|w_z\| = \sqrt{(-0.501232)^2 + (0.633958)^2} \ \approx\ 0.808, \qquad \boxed{\text{width} = \frac{2}{\|w_z\|}\ \approx\ 2.475}.$$

Эквивалентная форма через матрицу Грама:

$$\text{width} = \|Z_2 - Z_3\| = \sqrt{K_{22} - 2K_{23} + K_{33}} = \sqrt{6.12429712} \approx 2.475.$$

Возврат к исходному пространству x (подробно)

Связь $Z = (x-\mu)/\sigma$ даёт преобразование параметров:

$$\boxed{w_x = \frac{w_z}{\sigma}, \qquad b_x = b_z - w_z^\top \frac{\mu}{\sigma}}$$

Числа для нашего набора.

$$\begin{aligned}
\mu &= (0.537500,\ 0.175000),\\
\sigma &= (0.26061226,\ 0.17514280),\\
\frac{\mu}{\sigma} &= \left(\frac{0.5375}{0.26061226},\ \frac{0.175}{0.17514280}\right) \approx (2.062010,\ 0.999185).
\end{aligned}$$

$$\begin{aligned}
w_{x,1} &= \frac{-0.501232}{0.26061226} = -1.923287,\\
w_{x,2} &= \frac{0.633958}{0.17514280} = \quad 3.619663.
\end{aligned}$$

$$\begin{aligned}
w_z^\top \frac{\mu}{\sigma} &= (-0.501232)\cdot 2.062010 + (0.633958)\cdot 0.999185\\
&= -0.400326,
\end{aligned}$$

$$b_x = b_z - w_z^\top \frac{\mu}{\sigma} = -0.054025 - (-0.400326) = 0.346301.$$

Компоненты w_x:

$$w_{x,1} = \frac{-0.501232}{0.26061226} = -1.923287, \qquad w_{x,2} = \frac{0.633958}{0.17514280} = 3.619663,$$

скалярное произведение в поправке к b:

$$w_z^\top \frac{\mu}{\sigma} = (-0.501232) \cdot 2.062010 + (0.633958) \cdot 0.999185 \approx -0.400326,$$

поэтому

$$\boxed{b_x = b_z - w_z^\top \frac{\mu}{\sigma} = -0.054025 - (-0.400326) = 0.346301}.$$

Итоговая гиперплоскость в исходном пространстве x:

$$\boxed{w_x^\top x + b_x = 0 \quad \Longleftrightarrow \quad (-1.923287)\, x_1 + (3.619663)\, x_2 + 0.346301 = 0}.$$

Классификация: $\operatorname{sign}(w_x^\top x + b_x)$.

Проверка на исходных координатах: $w_x^\top x_i + b_x$ и знак

Напомним связь решений в Z и x:

$$z = \frac{x-\mu}{\sigma}, \qquad w_x = \frac{w_z}{\sigma}, \qquad b_x = b_z - w_z^\top \frac{\mu}{\sigma} \implies w_z^\top z + b_z \equiv w_x^\top x + b_x.$$

Численные параметры:

$$w_x = (-1.923287,\ 3.619663), \qquad b_x = 0.346301.$$

Таблица для всех точек:

i	$x_i = (x_{i1}, x_{i2})$	$Z_i = (z_{i1}, z_{i2})$	y_i	$w_x^\top x_i + b_x$	$\operatorname{sign}(w_x^\top x_i + b_x)$
1	(0.27, 0.36)	(−1.026429, 1.056281)	1	1.130092	1
2	(0.30, 0.34)	(−0.911316, 0.942088)	1	1.000000	1
3	(0.70, 0.00)	(0.623532, −0.999185)	−1	−1.000000	−1
4	(0.88, 0.00)	(1.314213, −0.999185)	−1	−1.346192	−1

Видно, что $\operatorname{sign}(w_x^\top x_i + b_x) = y_i$ для всех i.

Графики SVM

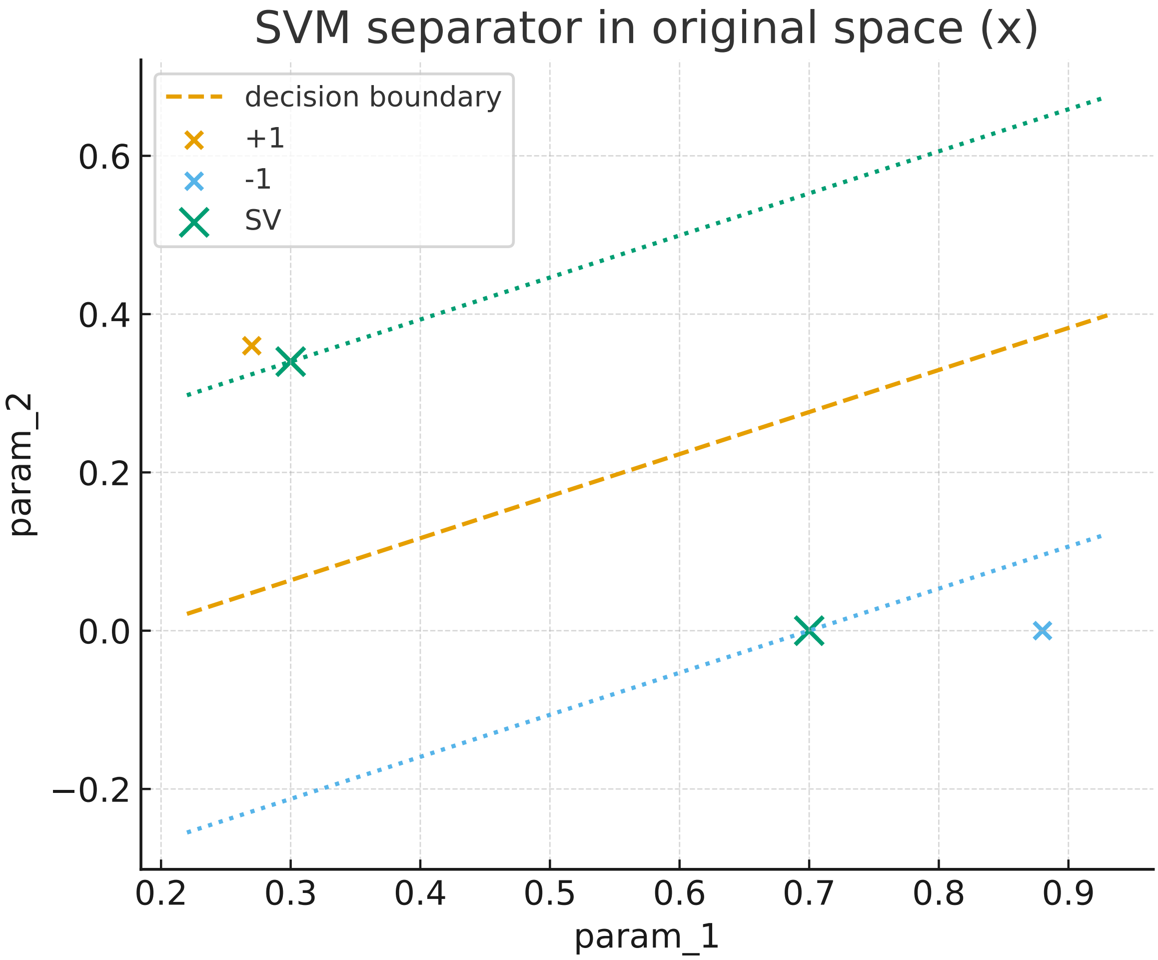

Рис. 1: Разделяющая прямая и полосы зазора в исходном пространстве x.

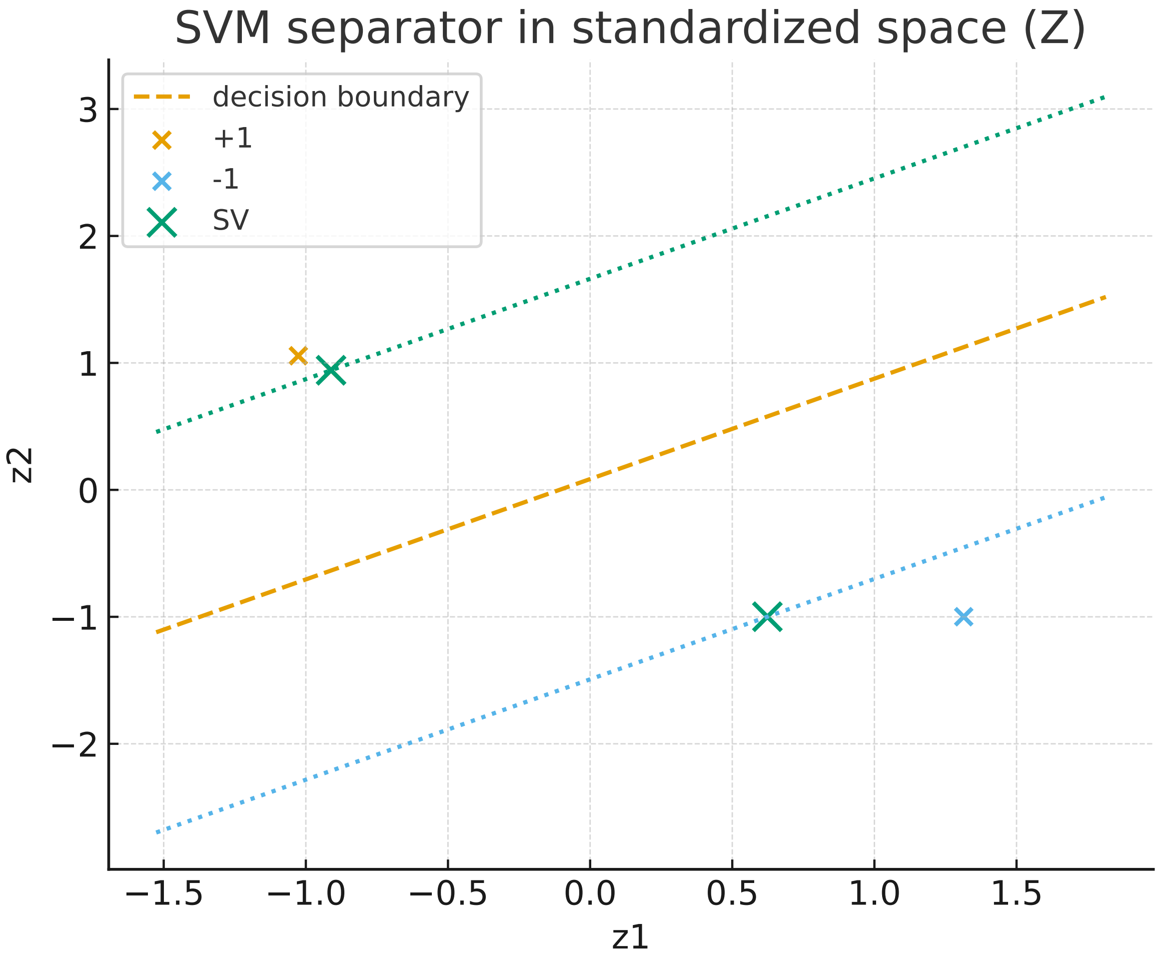

Рис. 2: Разделяющая прямая и полосы зазора в стандартизованном пространстве Z.