ANALÝZA A KLASIFIKACE DAT

prof. Ing. Jiří Holčík, CSc.

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

EVROPSKÁ UNIE

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ

II. PŘÍZNAKOVÁ KLASIFIKACE - ÚVOD

PŘÍZNAKOVÝ POPIS

Příznakový obraz x zpracovávaných dat je vyjádřen n-rozměrným (sloupcovým) vektorem hodnot x_i, $i=1,2,\ldots,n$ příznakových proměnných (veličin) charakterizujících vlastnosti těchto dat, tj. platí

$$\mathbf{x}=(x_1,x_2,\ldots,x_n)^T.$$

PŘÍZNAKOVÝ POPIS

Příznakové proměnné mohou popisovat kvantitativní i kvalitativní vlastnosti souboru dat. Jejich hodnoty nazýváme příznaky.

Podle definičního oboru rozlišujeme proměnné:

- spojité
- nespojité, diskrétní, vyjmenovatelné
- logické, binární, alternativní, dichotomické

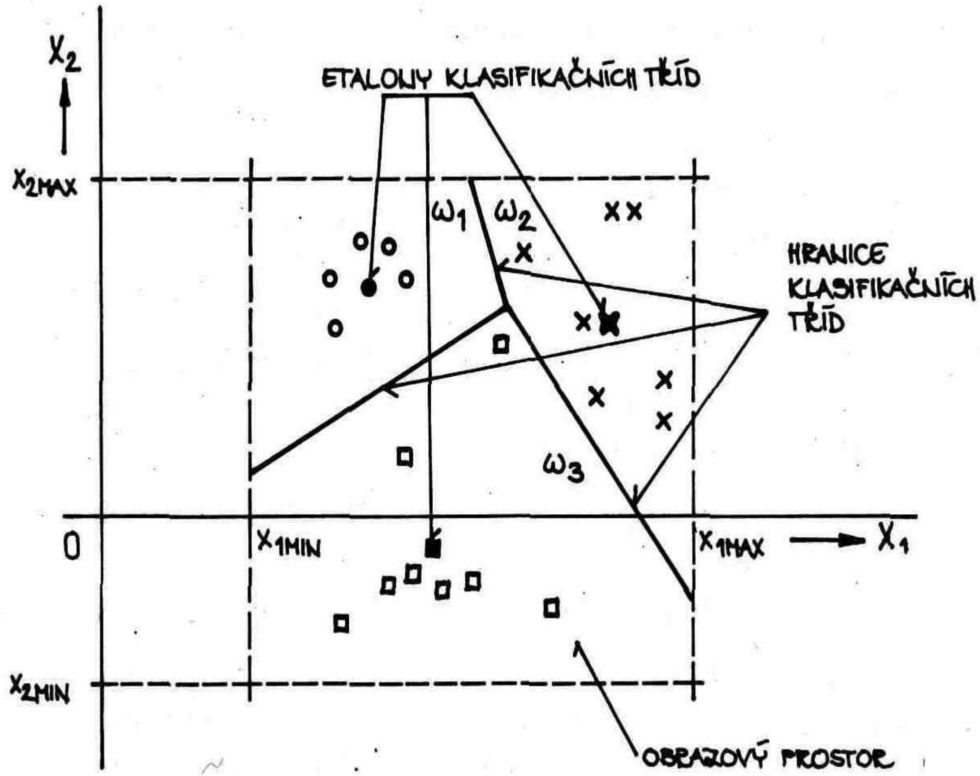
PŘÍZNAKOVÝ POPIS

Vrchol každého příznakového vektoru (obrazu) představuje bod n-rozměrného prostoru $\mathcal{X}^n$, který nazýváme obrazovým prostorem.

Obrazový prostor je definován kartézským součinem definičních oborů všech příznakovým proměnných, tzn. že jej tvoří všechny možné obrazy zpracovávaného souboru dat.

PŘÍZNAKOVÝ POPIS

Při vhodném výběru příznakových veličin je podobnost objektů (je popisujících dat) z jedné klasifikační třídy vyjádřena blízkostí jejich obrazů v obrazovém prostoru.



Vymezení klasifikační třídy:

- etalony - charakteristické reprezentativní obrazy
- hranice
- diskriminační funkce

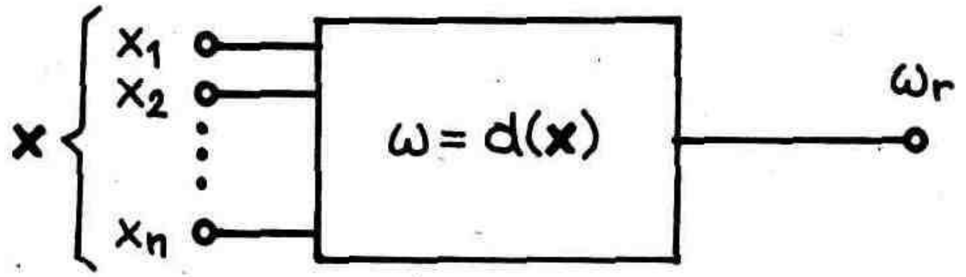
PŘÍZNAKOVÝ KLASIFIKÁTOR

Příznakový klasifikátor je stroj s tolika vstupy, kolik je příznaků a s jedním diskrétním výstupem, který udává třídu, do které klasifikátor zařadil rozpoznávaný obraz.

$$\omega_r = d(\mathbf{x})$$

$d(\mathbf{x})$ je skalární funkce vektorového argumentu $\mathbf{x}$, kterou nazýváme rozhodovací pravidlo klasifikátoru;

ω_r je identifikátor klasifikační třídy



PŘÍZNAKOVÝ KLASIFIKÁTOR

- ☑ deterministický a nedeterministický
- ☑ s pevným a proměnným počtem příznaků
- ☑ bez učení a s učením

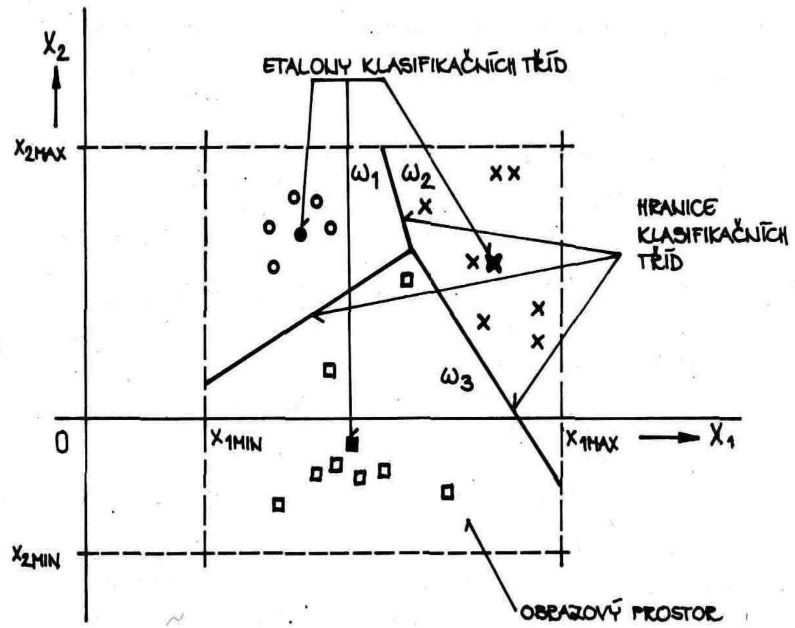
PŘÍZNAKOVÝ KLASIFIKÁTOR

- ☑ deterministický a nedeterministický
- ☑ s pevným a proměnným počtem příznaků
- ☑ bez učení a s učením

Nadále se nějaký čas věnujme deterministickým klasifikátorům s pevným počtem příznaků.

PŘÍZNAKOVÝ KLASIFIKÁTOR

- Obrazový prostor je rozhodovacím pravidlem rozdělen na R disjunktních prostorů $\mathcal{R}_r$, r=1,…,R, přičemž každá podmnožina $\mathcal{R}_r$ obsahuje ty obrazy **x**, pro které je $\omega_r = d(\mathbf{x})$.
- Návrh rozhodovacího pravidla je základním problémem návrhu klasifikátoru.



KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

DISKRIMINAČNÍ ANALÝZA

týká se obecně vztahu mezi kategoriální proměnnou a množinou vzájemně vázaných příznakových proměnných.

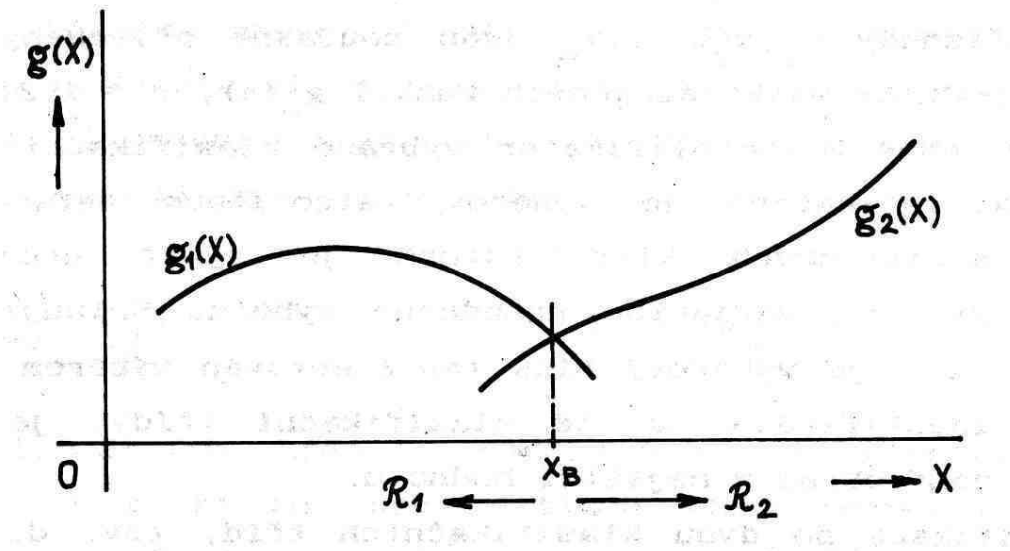
Konkrétně, předpokládejme že existuje konečný počet, řekněme R, různých a priori známých populací, kategorií, tříd nebo skupin, které označujeme ω_r, r=1,…,R a úkolem diskriminační analýzy je nalézt vztah, na základě kterého pro daný vektor příznaků popisujících konkrétní objekt tomuto vektoru přiřadíme hodnotu ω_r.

KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

- hranice klasifikačních tříd definujeme pomocí R skalárních funkcí $g_1(\mathbf{x})$, $g_2(\mathbf{x})$,…, $g_R(\mathbf{x})$ takových , že pro obraz **x** z podmnožiny $\mathcal{R}_r$ pro všechna r platí

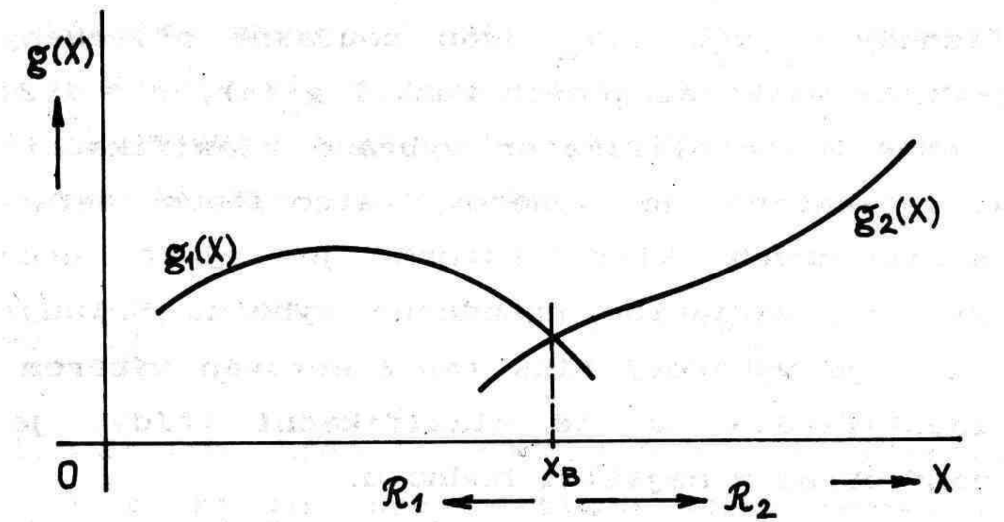
$$g_r(\mathbf{x}) > g_s(\mathbf{x}),\ \text{pro } s = 1,2,\ldots,R \text{ a } r \neq s$$

- funkce $g_r(\mathbf{x})$ mohou vyjadřovat např. míru výskytu obrazu **x** patřícího do r-té klasifikační třídy v daném místě obrazového prostoru – nazýváme je **diskriminační funkce**

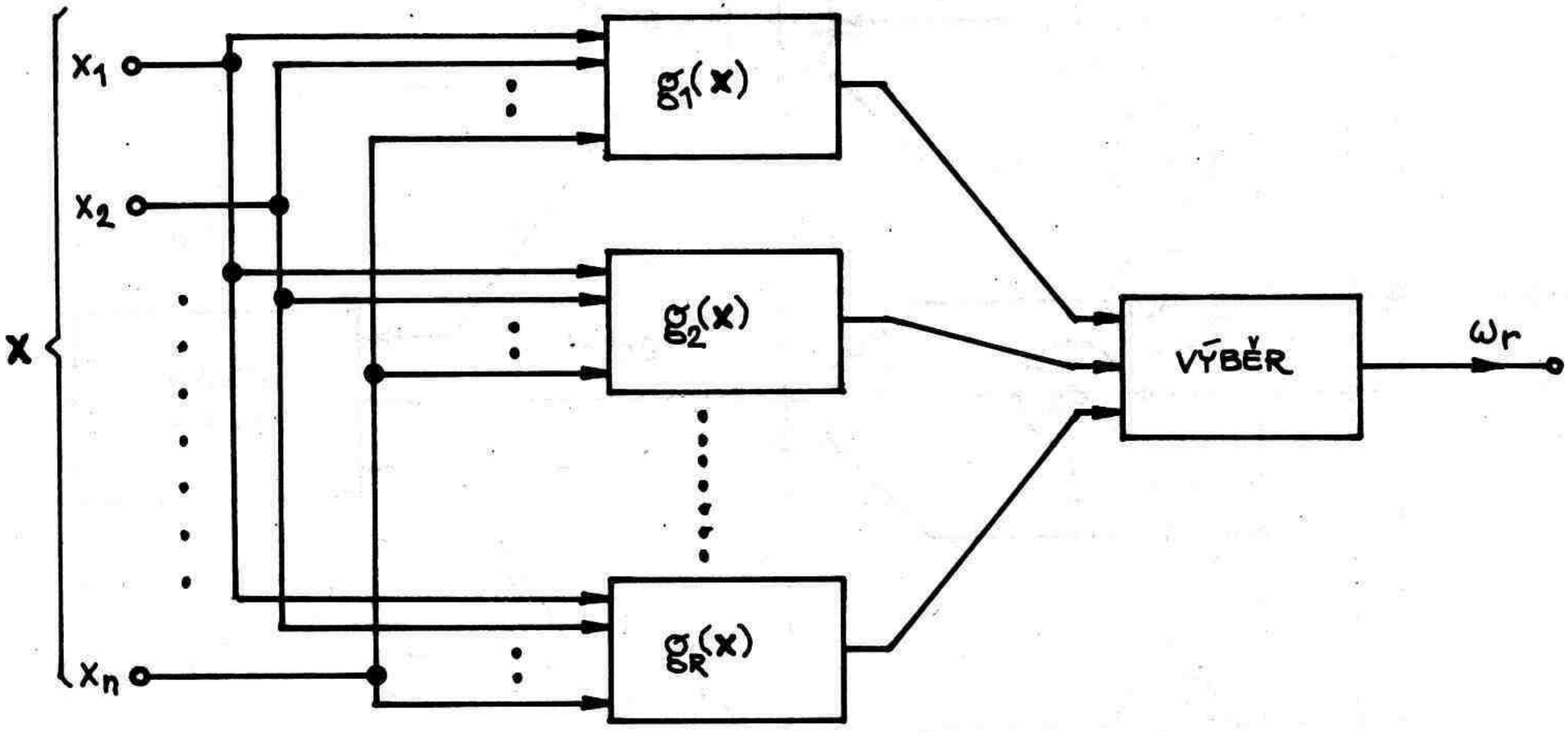


KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

- ☑ hranice mezi dvěma sousedními podmnožinami $\mathcal{R}_r$ a $\mathcal{R}_s$ je určena průmětem průsečíku funkcí $g_r(\mathbf{x})$ a $g_s(\mathbf{x})$, definovaného rovnicí $g_r(\mathbf{x}) = g_s(\mathbf{x})$, do obrazového prostoru.



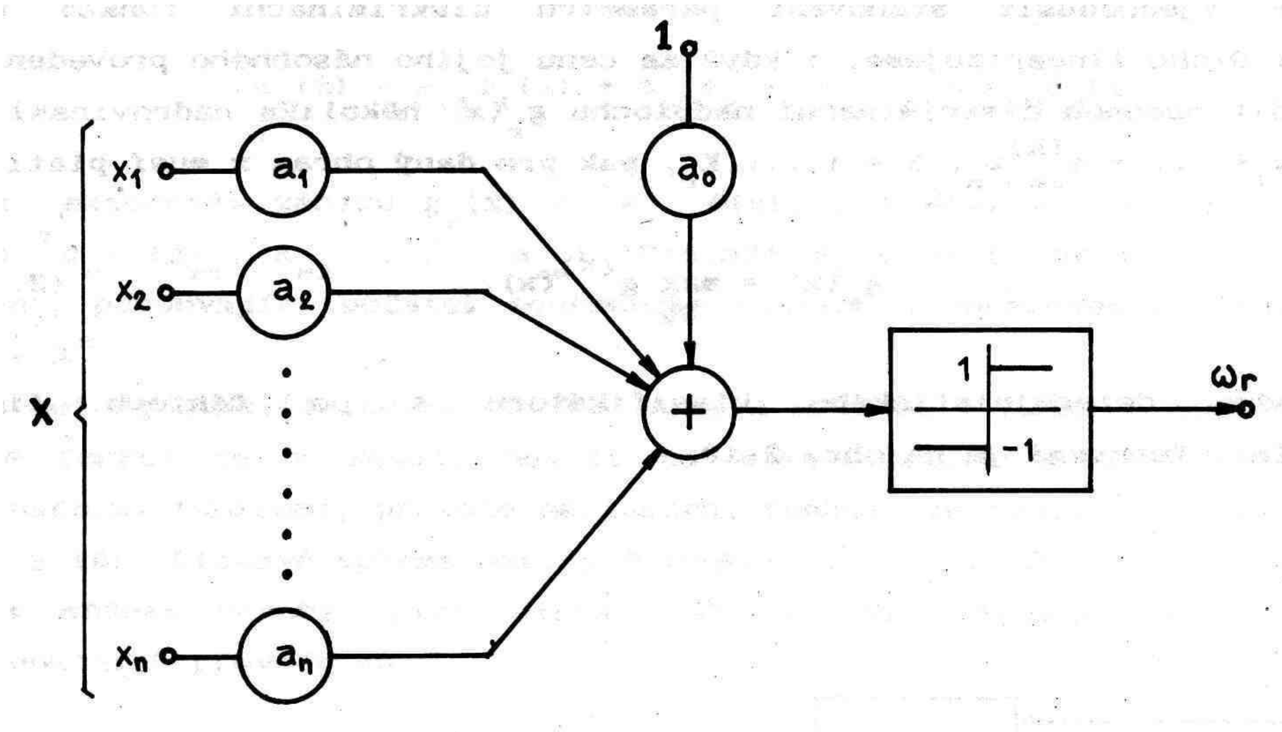
BLOKOVÉ SCHÉMA KLASIFIKÁTORU POMOCÍ DISKRIMINAČNÍCH FUNKCÍ



BLOKOVÉ SCHÉMA KLASIFIKÁTORU POMOCÍ DISKRIMINAČNÍCH FUNKCÍ

☑ u dichotomického klasifikátoru (dvě třídy) je

$$\omega = \mathrm{sign}\,(g_1(\mathbf{x}) - g_2(\mathbf{x}))$$



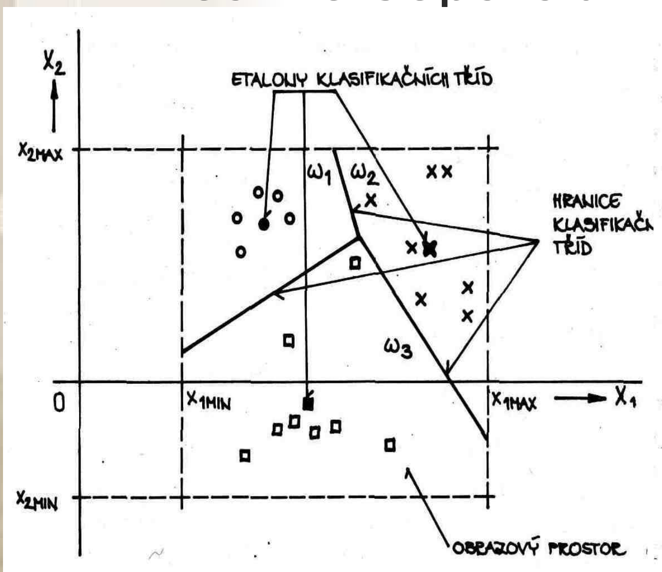
KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

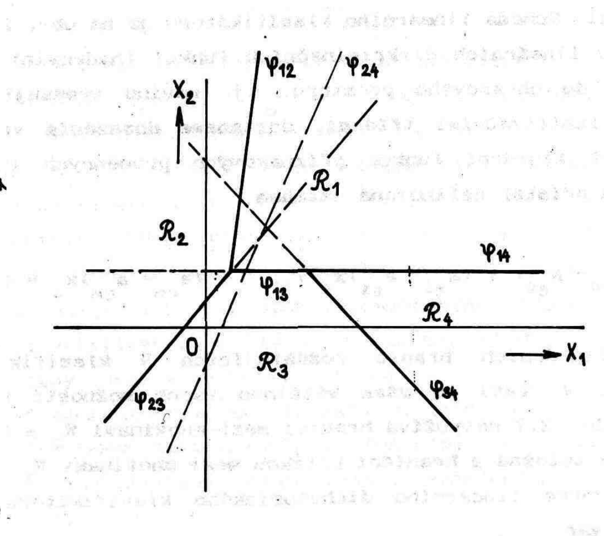
- nejjednodušším tvarem diskriminační funkce je funkce lineární, která má tvar

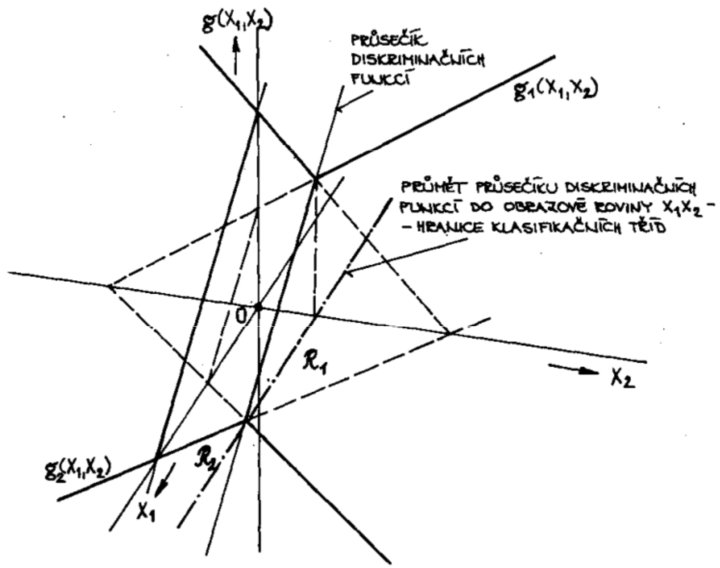
$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n$$

kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i-tého příznaku x_i

- lineárně separabilní třídy







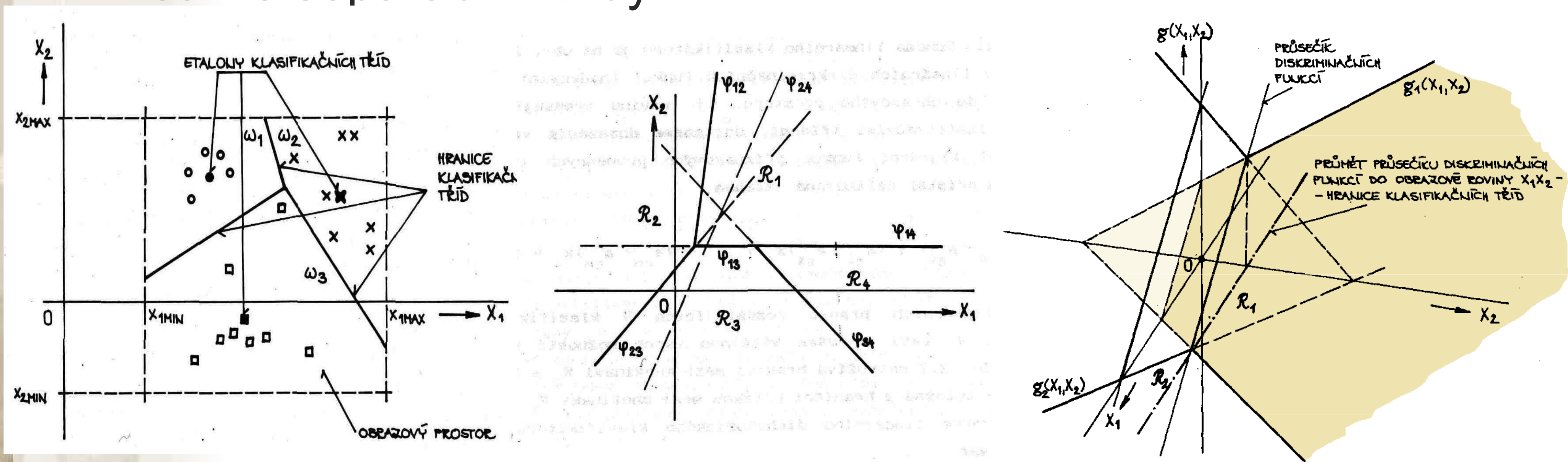
KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

- nejjednodušším tvarem diskriminační funkce je funkce lineární, která má tvar

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n$$

kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i-tého příznaku x_i

- lineárně separabilní třídy



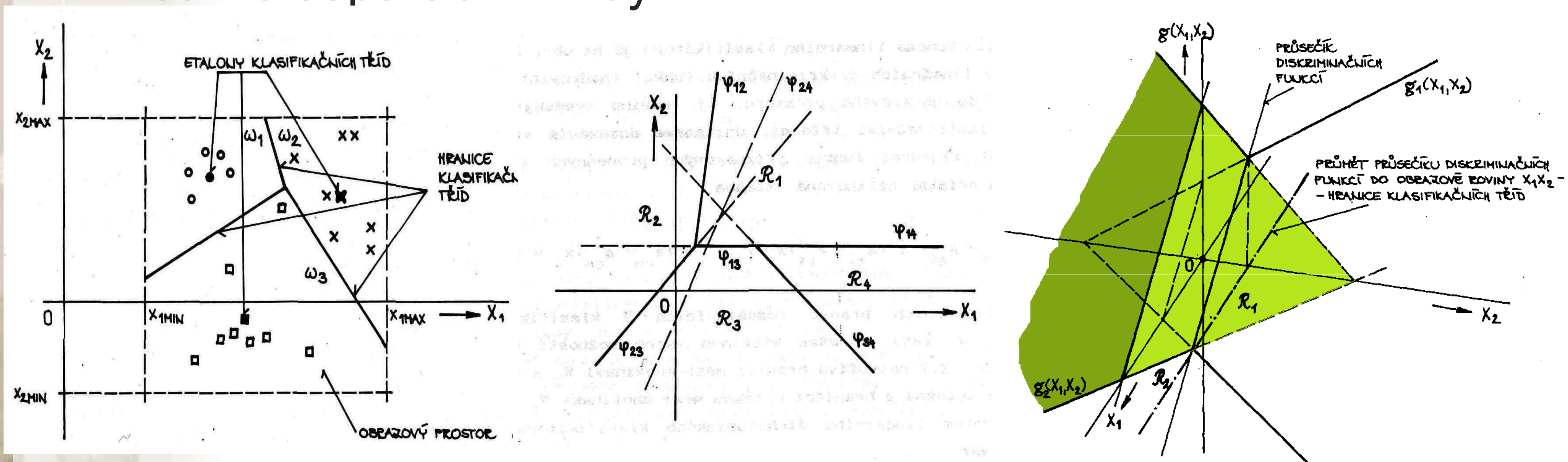
KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

- nejjednodušším tvarem diskriminační funkce je funkce lineární, která má tvar

$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n$$

kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i-tého příznaku x_i

- lineárně separabilní třídy



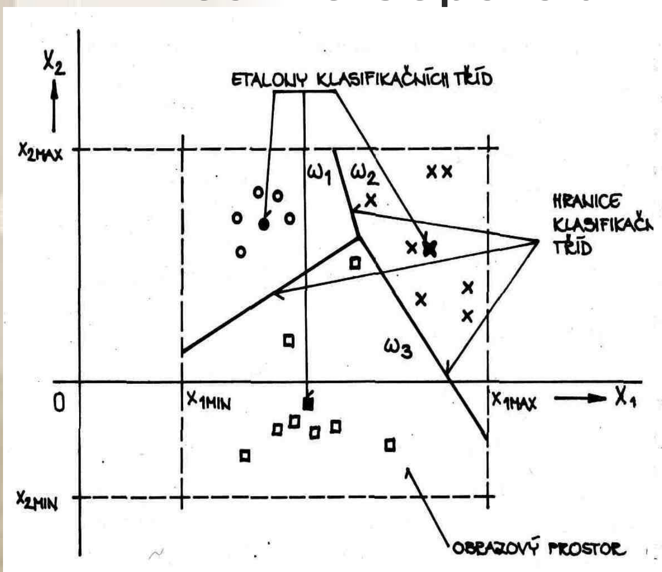
KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

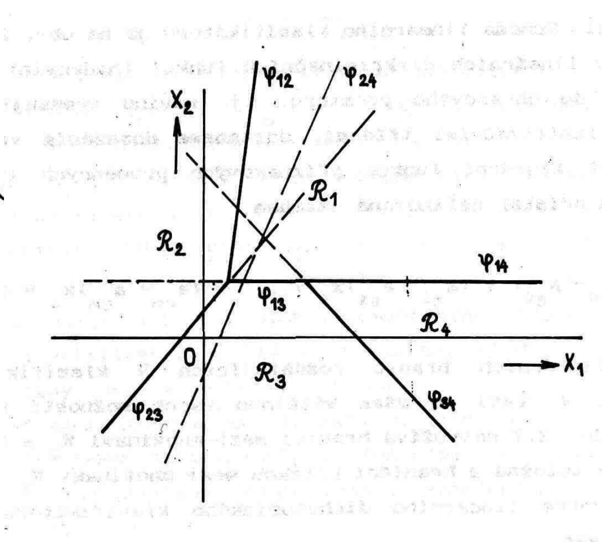
- nejjednodušším tvarem diskriminační funkce je funkce lineární, která má tvar

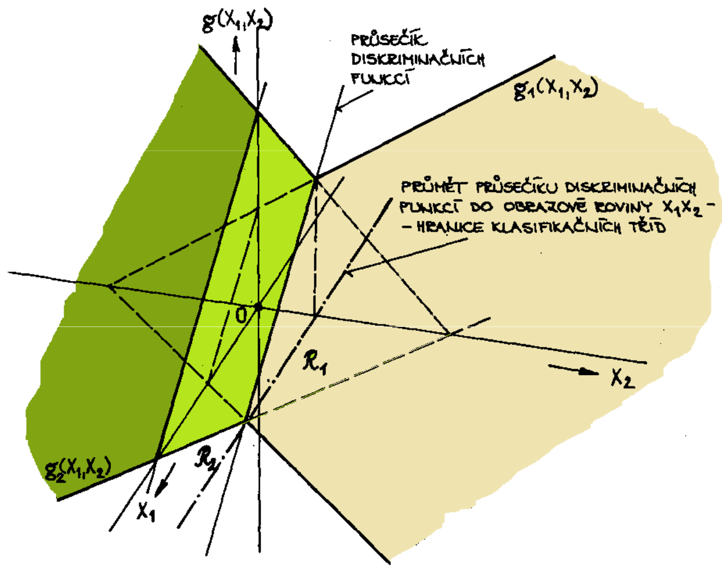
$$g_r(\mathbf{x}) = a_{r0} + a_{r1}x_1 + a_{r2}x_2 + \ldots + a_{rn}x_n$$

kde a_{r0} je práh diskriminační funkce posouvající počátek souřadného systému a a_{ri} jsou váhové koeficienty i-tého příznaku x_i

- lineárně separabilní třídy



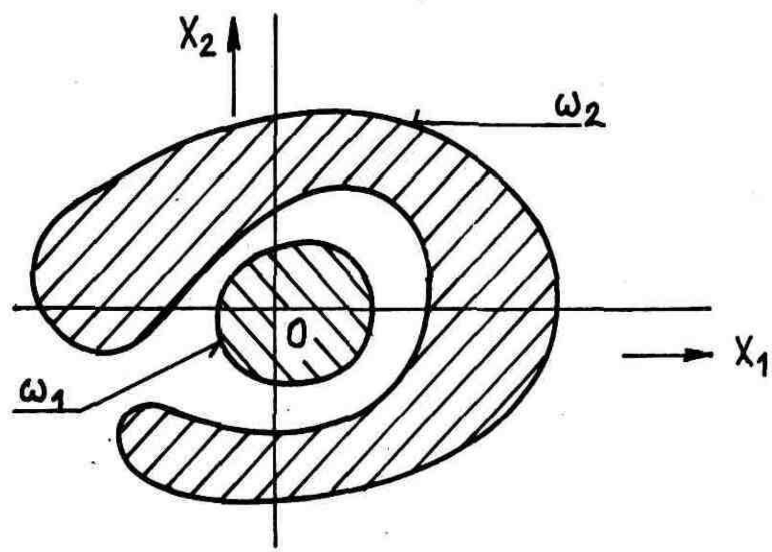




KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

LINEÁRNĚ NESEPARABILNÍ TŘÍDY

- zachováme původní obrazový prostor a zvolíme nelineární diskriminační funkci
  - definovanou obecně
  - složenou po částech z lineárních úseků



- zobrazíme původní n-rozměrný obrazový prostor X^n nelineární transformací Φ: $X^n \rightarrow X^m$ do nového m-rozměrného prostoru X^m, obecně je $m \neq n$, tak, aby v novém prostoru byly klasifikační třídy lineárně separabilní a v novém prostoru použijeme lineární klasifikátor (Φ převodník)

KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

BAYESŮV KLASIFIKÁTOR

KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

BAYESŮV KLASIFIKÁTOR

při řešení praktických úloh je třeba předpokládat, že obrazy signálů jsou ovlivněny víceméně náhodnými fluktuacemi zdroje signálu, v přenosové cestě, při předzpracování i analýze, které se nepodaří zcela eliminovat.

KLASIFIKACE PODLE DISKRIMINAČNÍCH FUNKCÍ

BAYESŮV KLASIFIKÁTOR

$$P(\omega_r|\mathbf{x}) = \frac{p(\mathbf{x}|\omega_r).P(\omega_r)}{p(\mathbf{x})}$$

$P(\omega_r|\mathbf{x})$ je aposteriorní podmíněná pravděpodobnost zatřídění obrazového vektoru **x** do třídy ω_r;

$p(\mathbf{x}|\omega_r)$ je podmíněná hustota pravděpodobnosti obrazů **x** ve třídě ω_r;

$P(\omega_r)$ je apriorní pravděpodobnost třídy ω_r;

$p(\mathbf{x})$ je hustota pravděpodobnosti rozložení všech obrazů **x** v celém obrazovém prostoru.

ZÁKLADNÍ POJMY A PŘEDPOKLADY

- ☑ ztrátová funkce $\lambda(\omega_r|\omega_s)$ udává ztrátu při chybné klasifikaci obrazu ze třídy ω_s do třídy ω_r.
- ☑ matice ztrátových funkcí

$$\boldsymbol{\lambda} = \begin{bmatrix} \lambda(\omega_1|\omega_1) & \lambda(\omega_1|\omega_2) & \cdots & \lambda(\omega_1|\omega_R) \\ \lambda(\omega_2|\omega_1) & \lambda(\omega_2|\omega_2) & \cdots & \lambda(\omega_2|\omega_R) \\ \vdots & \vdots & \ddots & \vdots \\ \lambda(\omega_R|\omega_1) & \lambda(\omega_R|\omega_2) & \cdots & \lambda(\omega_R|\omega_R) \end{bmatrix}$$

- ☑ střední ztráta $J(\mathbf{a})$ udává průměrnou ztrátu při chybné klasifikaci obrazu $\mathbf{x}$

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- pokud se soustředíme na obrazy pouze ze třídy ω_s, je střední ztráta dána průměrnou hodnotou z $\lambda(d(\mathbf{x},\mathbf{a})|\omega_s)$ vzhledem ke všem obrazům ze třídy ω_s, tj.

$$J_s(\mathbf{a}) = \int_x \lambda(d(\mathbf{x},\mathbf{a})|\omega_s).p(\mathbf{x}|\omega_s)d\mathbf{x}$$

kde $p(\mathbf{x}|\omega_s)$ je podmíněná hustota pravděpodobnosti výskytu obrazu **x** ve třídě ω_s

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- Celková střední ztráta J(**a**) je průměrná hodnota ze ztrát $J_s(\mathbf{a})$

$$J(\mathbf{a}) = \sum_{s=1}^{R} J_s(\mathbf{a}).P(\omega_s) = \int_{\mathcal{X}} \sum_{s=1}^{R} \lambda(d(\mathbf{x},\mathbf{a})|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x}$$

- nebo podle Bayesova vzorce ($P(\omega_s|\mathbf{x}).p(\mathbf{x}) = p(\mathbf{x}|\omega_s).P(\omega_s)$)

$$J(\mathbf{a}) = \int_{\mathcal{X}} \sum_{s=1}^{R} \lambda(d(\mathbf{x},\mathbf{a})|\omega_s).p(\mathbf{x}).P(\omega_s|\mathbf{x})d\mathbf{x}$$

kde p(**x**) je hustota pravděpodobnosti výskytu obrazu **x** v celém obrazovém prostoru a $P(\omega_s|\mathbf{x})$ je podmíněná pravděpodobnost, že daný obraz patří do třídy ω_s (tzv. aposteriorní pravděpodobnost třídy ω_s.

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- Návrh optimálního klasifikátoru, který by minimalizoval střední ztrátu, spočívá v nalezení takové množiny parametrů rozhodovacího pravidla **a***, že platí

$$J(\mathbf{a}^*) = \min_{\mathbf{a}} J(\mathbf{a})$$

- Dosadíme-li za J(**a**) z předchozího vztahu, je

$$J(\mathbf{a}^*) = \min_{\mathbf{a}} \int_{\mathcal{X}} \sum_{s=1}^{R} \lambda(d(\mathbf{x},\mathbf{a})|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x}$$

- Je-li ztrátová funkce $\lambda(\omega_r|\omega_s)$ konstantní pro všechny obrazy z ω_s, je dále

$$J(\mathbf{a}^*) = \int_{\mathcal{X}} \min_{r} \sum_{s=1}^{R} \lambda(\omega_r|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x}$$

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- Označíme-li ztrátu při klasifikaci obrazu **x** do třídy ω_r

$$L_x(\omega_r) = \sum_{s=1}^{R} \lambda(\omega_r | \omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)$$

tak po dosazení dostaneme

$$J(\mathbf{a}^*) = \int_{\mathcal{X}} \min_r L_x(\omega_r) d\mathbf{x}$$

Úloha nalezení minima celkové střední ztráty se tak převedla na minimalizaci funkce $L_x(\omega_r)$. Optimální rozhodovací pravidlo $d(\mathbf{x},\mathbf{a}^*)$ podle kritéria minimální celkové střední ztráty je

$$L_{\mathbf{x}}\left(d_{ME}(\mathbf{x},\mathbf{a}^*)\right) = \min_r L_x(\omega_r)$$

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY

- Chceme-li využít principu diskriminačních funkcí

$$\min L_x(\omega_r) = \max\left(-L_x(\omega_r)\right)$$

- Diskriminační funkci optimálního klasifikátoru podle kritéria minimální chyby pak definujeme

$$g_r(\mathbf{x}) = -L_x(\omega_r) = -\sum_{s=1}^{R} \lambda(\omega_r|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)$$

KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY
DICHOTOMICKÝ KLASIFIKÁTOR

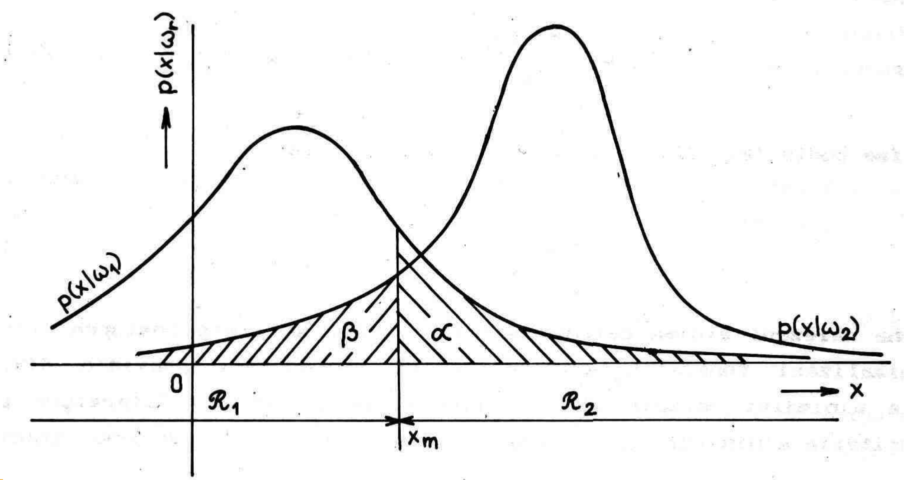
Celková střední ztráta v případě dvou tříd je

$$J(\mathbf{a}) = \int_{\mathcal{R}_1} \sum_{s=1}^{2} \lambda(\omega_1|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x} + \int_{\mathcal{R}_2} \sum_{s=1}^{2} \lambda(\omega_2|\omega_s).p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x} =$$

$$= \lambda(\omega_1|\omega_1).P(\omega_1)\int_{\mathcal{R}_1} p(\mathbf{x}|\omega_1).d\mathbf{x} + \lambda(\omega_1|\omega_2).P(\omega_2)\int_{\mathcal{R}_1} p(\mathbf{x}|\omega_2).d\mathbf{x} +$$

$$+ \lambda(\omega_2|\omega_1).P(\omega_1)\int_{\mathcal{R}_2} p(\mathbf{x}|\omega_1).d\mathbf{x} + \lambda(\omega_2|\omega_2).P(\omega_2)\int_{\mathcal{R}_2} p(\mathbf{x}|\omega_2).d\mathbf{x} =$$

$$= \lambda(\omega_1|\omega_1).P(\omega_1).(1-\alpha) + \lambda(\omega_1|\omega_2).P(\omega_2).\beta + \lambda(\omega_2|\omega_1).P(\omega_1).\alpha + \lambda(\omega_2|\omega_2).P(\omega_2)(1-\beta)$$



KRITÉRIUM MINIMÁLNÍ STŘEDNÍ ZTRÁTY
DICHOTOMICKÝ KLASIFIKÁTOR

Diskriminační funkce pro dichotomický klasifikátor bude

$$g(\mathbf{x}) = g_1(\mathbf{x}) - g_2(\mathbf{x}) = -L_{\mathbf{x}}(\omega_1) + L_{\mathbf{x}}(\omega_2) =$$

$$= -\lambda(\omega_1|\omega_1).p(\mathbf{x}|\omega_1).P(\omega_1) - \lambda(\omega_1|\omega_2).p(\mathbf{x}|\omega_2).P(\omega_2) +$$

$$+ \lambda(\omega_2|\omega_1).p(\mathbf{x}|\omega_1).P(\omega_1) + \lambda(\omega_2|\omega_2).p(\mathbf{x}|\omega_2).P(\omega_2) =$$

$$= \left(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)\right).p(\mathbf{x}|\omega_1).P(\omega_1) + \left(\lambda(\omega_2|\omega_2) - \lambda(\omega_1|\omega_2)\right).p(\mathbf{x}|\omega_2).P(\omega_2)$$

Položíme-li tento výraz nule dostaneme vztah pro hraniční plochu dichotomického klasifikátoru, ze kterého můžeme určit poměr hustot pravděpodobnosti výskytu obrazu **x** v každé z obou klasifikačních tříd - **věrohodnostní poměr**

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{\left(\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2)\right).P(\omega_2)}{\left(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)\right).P(\omega_1)}$$

Obraz **x** zařadíme do třídy ω_1, když je věrohodnostní poměr větší než výraz na pravé straně, je-li menší pak obraz **x** zařadíme do třídy ω_2.

VĚROHODNOSTNÍ POMĚR I.

- Sumarizuje veškerou informaci získanou experimentem.
- Pravděpodobnost, že jev (data) nastane za daných podmínek (hypotéza) děleno pravděpodobností, že stejný jev nastane za jiných podmínek. Podmínky jsou vzájemně se vylučující.

VĚROHODNOSTNÍ POMĚR II.

Věrohodnostní poměr (likelihood ratio) LR udává podíl pravděpodobnosti, že se vyskytne nějaký jev *A* za určité podmínky (jev *B*), k pravděpodobnosti, že se jev *A* vyskytne, když podmínka neplatí (jev nonB). Má-li například pacient náhlou ztrátu paměti (jev *A*), chceme znát věrohodnostní poměr výskytu jevu *A* v případě, že má mozkový nádor (jev *B*), tj. podíl pravděpodobnosti, s jakou ztráta paměti vzniká při nádoru mozku, k pravděpodobnosti, s jakou vzniká v ostatních případech . Věrohodnostní poměr je tedy podíl podmíněných pravděpodobností

$$LR = \frac{P(A|B)}{P(A|nonB)}$$

KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOSTI CHYBNÉHO ROZHODNUTÍ

Díky obtížnému stanovení hodnot ztrátových funkcí $\lambda(\omega_r|\omega_s)$ se kritérium minimální chyby zjednodušuje použitím jednotkových ztrátových funkcí definovaných

$$\lambda(\omega_r|\omega_s) = \begin{cases} 0 & \text{pro } r = s \\ 1 & \text{pro } r \neq s \end{cases}$$

Matice jednotkových ztrátových funkcí má pak tvar

$$\boldsymbol{\lambda} = \begin{bmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 0 \end{bmatrix}$$

a celková ztráta je

$$J(\mathbf{a}) = \sum_{\substack{s=1 \\ s\neq r}}^{R} \int_{\mathcal{X}-\mathcal{R}_s} p(\mathbf{x}|\omega_s).P(\omega_s)d\mathbf{x}$$

což je hodnota pravděpodobnosti chybného rozhodnutí.

KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOSTI CHYBNÉHO ROZHODNUTÍ

Dosadíme-li hodnoty jednotkových ztrátových funkcí do vztahu pro ztrátu při klasifikaci obrazu do chybné třídy

$$L_{\mathbf{x}}(\omega_r) = \sum_{\substack{s=1 \\ s\neq r}}^{R} p(\mathbf{x}|\omega_r).P(\omega_s) = \sum_{s=1}^{R} p(\mathbf{x}|\omega_s).P(\omega_s) - p(\mathbf{x}|\omega_r).P(\omega_r)$$

a s využitím Bayesova vztahu

$$L_{\mathbf{x}}(\omega_r) = p(\mathbf{x})\sum_{s=1}^{R} P(\omega_s|\mathbf{x}) - p(\mathbf{x}|\omega_r).P(\omega_r) = p(\mathbf{x}) - p(\mathbf{x}|\omega_r).P(\omega_r)$$

$p(\mathbf{x})$ nezávisí na klasifikační třídě a tedy neovlivňuje výběr minima.

Diskriminační funkci tedy můžeme určit jako

$$g(\mathbf{x}) = p(\mathbf{x}|\omega_r).P(\omega_r)$$

KRITÉRIUM MINIMÁLNÍ PRAVDĚPODOBNOSTI CHYBNÉHO ROZHODNUTÍ

V případě dichotomického klasifikátoru je diskriminační funkce

$$g(\mathbf{x}) = p(\mathbf{x}|\omega_1).P(\omega_1) - p(\mathbf{x}|\omega_2).P(\omega_2)$$

A věrohodnostní poměr je potom

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{P(\omega_2)}{P(\omega_1)}$$

KRITÉRIUM MAXIMÁLNÍ APOSTERIORNÍ PRAVDĚPODOBNOSTI

- Modifikujeme-li vztah pro ztrátu při chybné klasifikaci obrazu podle Bayesova vztahu ($P(\omega_s|\mathbf{x}).p(\mathbf{x}) = p(\mathbf{x}|\omega_s).P(\omega_s)$) platí

$$L_x(\omega_r) = \sum_{s=1}^{R} \lambda(\omega_r|\omega_s).p(\mathbf{x}).P(\omega_s|\mathbf{x}) = p(\mathbf{x})\sum_{s=1}^{R} \lambda(\omega_r|\omega_s).P(\omega_s|\mathbf{x})$$

- Hustota pravděpodobnosti $p(\mathbf{x})$ nezávisí na klasifikační třídě a tedy místo $L_x(\omega_r)$ lze použít

$$L'_x(\omega_r) = \frac{L_x(\omega_r)}{p(\mathbf{x})} = \sum_{s=1}^{R} \lambda(\omega_r|\omega_s).P(\omega_s|\mathbf{x})$$

a s jednotkovými ztrátovými funkcemi je

$$L'_x(\omega_r) = \sum_{\substack{s=1 \\ s\neq r}}^{R} P(\omega_s|\mathbf{x}) = \sum_{s=1}^{R} P(\omega_s|\mathbf{x}) - P(\omega_r|\mathbf{x}) = 1 - P(\omega_r|\mathbf{x})$$

KRITÉRIUM MAXIMÁLNÍ APOSTERIORNÍ PRAVDĚPODOBNOSTI

- Minimum ztráty $L'_x(\omega_r)$ je právě tehdy, když $P(\omega_r|\mathbf{x})$ je maximální. Tzn. že jako diskriminační funkci můžeme zvolit právě hodnotu aposteriorní pravděpodobnosti třídy ω_r, tj.

$$g_r(\mathbf{x}) = P(\omega_r|\mathbf{x})$$

- Pro případ dichotomického klasifikátoru je diskriminační funkce

$$g(\mathbf{x}) = P(\omega_1|\mathbf{x}) - P(\omega_2|\mathbf{x}) = 0.$$

Z toho plyne, že hranicí mezi třídami určuje vztah

$$P(\omega_1|\mathbf{x}) = P(\omega_2|\mathbf{x})$$

nebo

$$\frac{P(\omega_1|\mathbf{x})}{P(\omega_2|\mathbf{x})} = 1$$

Podle tohoto kritéria zatřídíme obraz do té třídy, jejíž aposteriorní pravděpodobnost je při výskytu obrazu **x** větší.

KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOSTI (MINIMAX)

Neznáme-li apriorní pravděpodobnosti všech tříd, předpokládáme rovnoměrné rozložení (pravděpodobnost všech tříd je táž ($P(\omega_s)$ = = $P(\omega)$ =1/R). Potom celková střední ztráta

$$J(\mathbf{a}) = \frac{1}{R}\sum_{s=1}^{R}\int_{\mathcal{X}} \lambda(\omega_r|\omega_s).p(\mathbf{x}|\omega_s)d\mathbf{x}$$

dosáhne minima, když

$$J(\mathbf{a}^*) = \frac{1}{R}\min_{\forall \mathbf{a}}\int_{\mathcal{X}}\sum_{s=1}^{R} \lambda(\omega_r|\omega_s).p(\mathbf{x}|\omega_s)d\mathbf{x}$$

Diskriminační funkci lze jako v předchozích případech definovat jako

$$g_r(\mathbf{x}) = -L_x(\omega_r) = -\sum_{s=1}^{R} \lambda(\omega_r|\omega_s).p(\mathbf{x}|\omega_s)$$

KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOSTI (MINIMAX)

- V případě dichotomie je věrohodnostní poměr

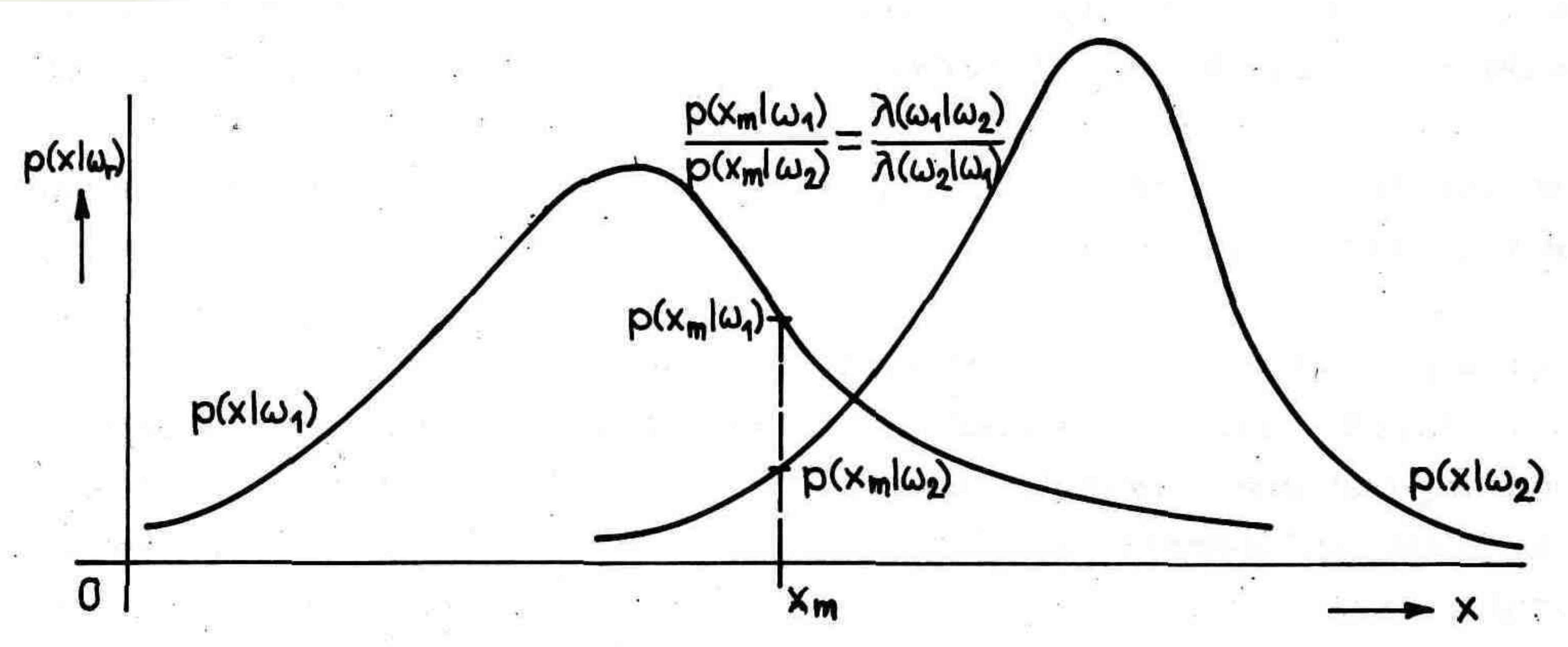
$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{\left(\lambda(\omega_1|\omega_2) - \lambda(\omega_2|\omega_2)\right)}{\left(\lambda(\omega_2|\omega_1) - \lambda(\omega_1|\omega_1)\right)}$$

- Pokud jsou ceny správného rozhodnutí nulové, tj. $\lambda(\omega_1|\omega_1) = \lambda(\omega_2|\omega_2) = 0$, je

$$\Lambda_{12} = \frac{p(\mathbf{x}|\omega_1)}{p(\mathbf{x}|\omega_2)} = \frac{\left(\lambda(\omega_1|\omega_2)\right)}{\left(\lambda(\omega_2|\omega_1)\right)}$$

- Obraz je zařazen do třídy ω_1, když je věrohodnostní poměr než poměr cen ztrát chybných zatřídění. Jsou-li obě ceny stejné, je obraz zařazen do té třídy, pro kterou je hodnota $p(\mathbf{x}|\omega_s)$ větší.

KRITÉRIUM MAXIMÁLNÍ PRAVDĚPODOBNOSTI (MINIMAX)



Příprava nových učebních materiálů pro obor Matematická biologie

je podporována projektem ESF

č. CZ.1.07/2.2.00/07.0318

„VÍCEOBOROVÁ INOVACE STUDIA MATEMATICKÉ BIOLOGIE"

EVROPSKÁ UNIE

MINISTERSTVO ŠKOLSTVÍ, MLÁDEŽE A TĚLOVÝCHOVY

OP Vzdělávání pro konkurenceschopnost

INVESTICE DO ROZVOJE VZDĚLÁVÁNÍ