

Facoltà: : BioScienze e Tecnologie Agro-Alimentari e Ambientali

Denominazione Corso di Laurea: Biotecnologie Avanzate (Laurea Magistrale)

Corso: Statistica e bioinformatica per le biotecnologie

MODULO:

Chemometria applicata (5 CFU, 40 ore)

Docente: Marcello Mascini

(mmascini@unite.it)

Il Docente e' disponibile per chiarimenti al termine della lezione o su richiesta via mail

3UD Analisi multivariata “supervised” (2 CFU = 16 ore)

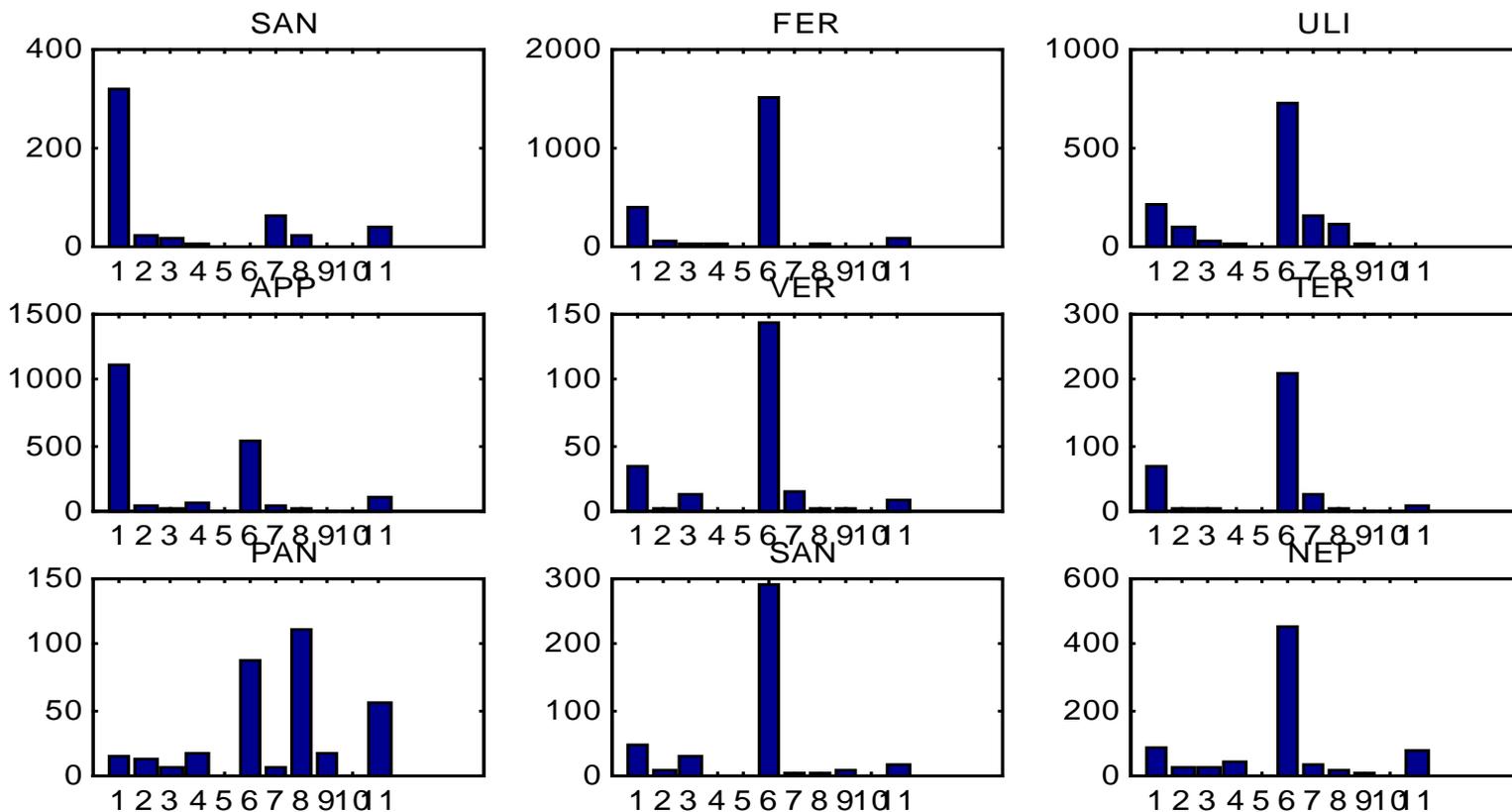
Differenze tra “unsupervised” e “supervised” analysis; Partial least square regression (PLS); Ordinary Least Squares regression (OLS); PLS discriminant analysis (PLS-DA); Esempi di elaborazione dati “supervised”.

Definizioni

- Definizioni fondamentali:
 - Pattern: serie di descrittori (features) che definiscono le proprietà di un soggetto complesso.
 - Classe: insieme a cui un soggetto appartiene.
 - Classificare: operazione matematica per cui un campione, descritto da una serie di features, viene assegnato ad una particolare classe.
 - L'insieme delle features è detto Pattern, l'operazione di classificazione si dice Pattern Recognition.
- La relazione tra campione e classe non è univoca, dipende dalle features scelte per descrivere il soggetto
- La pattern recognition è un problema "interessante" quando le singole features non sono in grado di identificare i campioni che descrivono
- Esempio:
 - Frutti:
 - Features: peso, forma, colore, contenuto di zuccheri, di acidi,.....
 - Alcuni di questi descrittori consentono di classificare per specie, altri per grado di maturità,...

Esempio di pattern acque minerali

- Profilo ionico di acque minerali



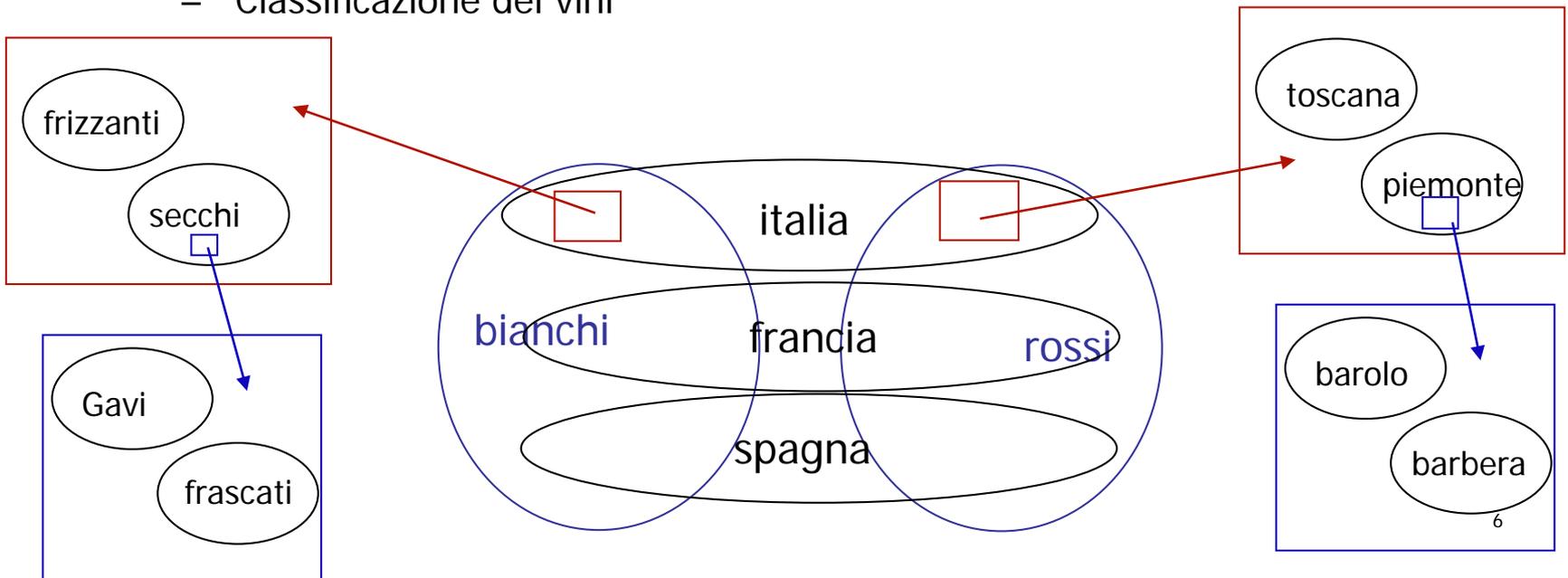
Ca	Na	Mg
K	NH4	HCO
SO4	Cl	NO3
F	SiO	

Pattern analysis

- La pattern recognition consiste nel estrarre informazioni rispetto a campioni descritti da patterns
- Matematicamente un pattern è un vettore i cui elementi corrispondono a "caratteristiche" (features) che descrivono aspetti parziali di un campione (o fenomeno) complesso
- La scelta delle features è legata al tipo di informazione che vuole essere descritta
- L'operazione ultima della pattern recognition è quella di assegnare un campione ad una classe di appartenenza (membership class)

Membership class

- Le classi di membership sono insiemi teorici i cui elementi condividono una caratteristica globale
- Gli elementi possono essere raggruppati secondo differenti schemi di classificazione a seconda delle caratteristiche globali ritenute interessanti.
- Esempio:
 - Classificazione dei vini

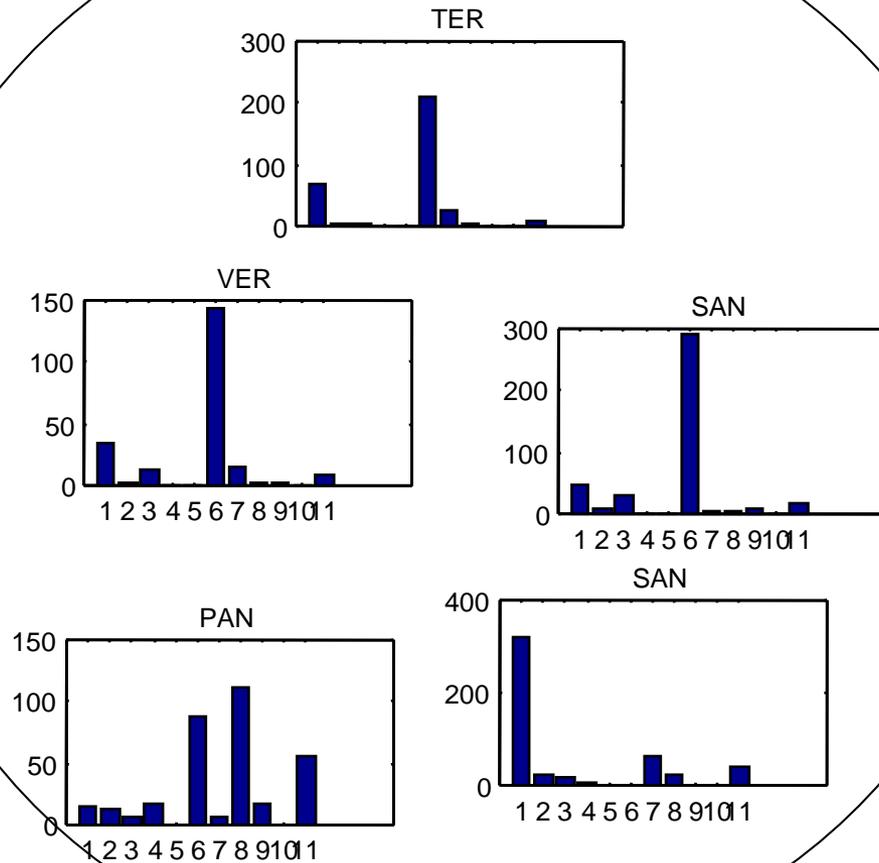


Patterns e membership class

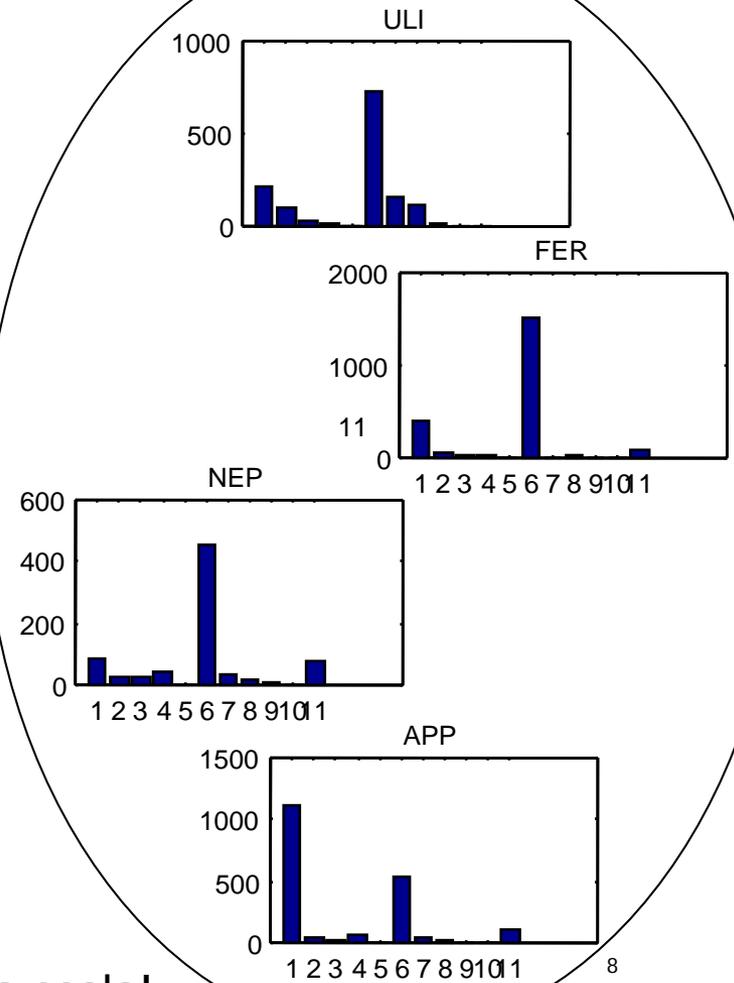
- Se le features sono appropriate per la classificazione voluta allora gli elementi della stessa classe avranno pattern simili.
- La similitudine tra pattern può essere evidenziata con una ispezione visuale di grafici opportuni
- I più semplici: istogramma e radar-plot
 - Esempio acque minerali:

istogrammi e membership class

oligominerali



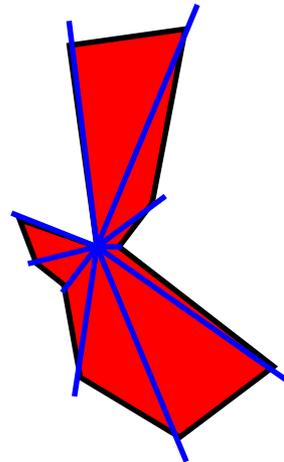
minerali



attenzione alla scala!

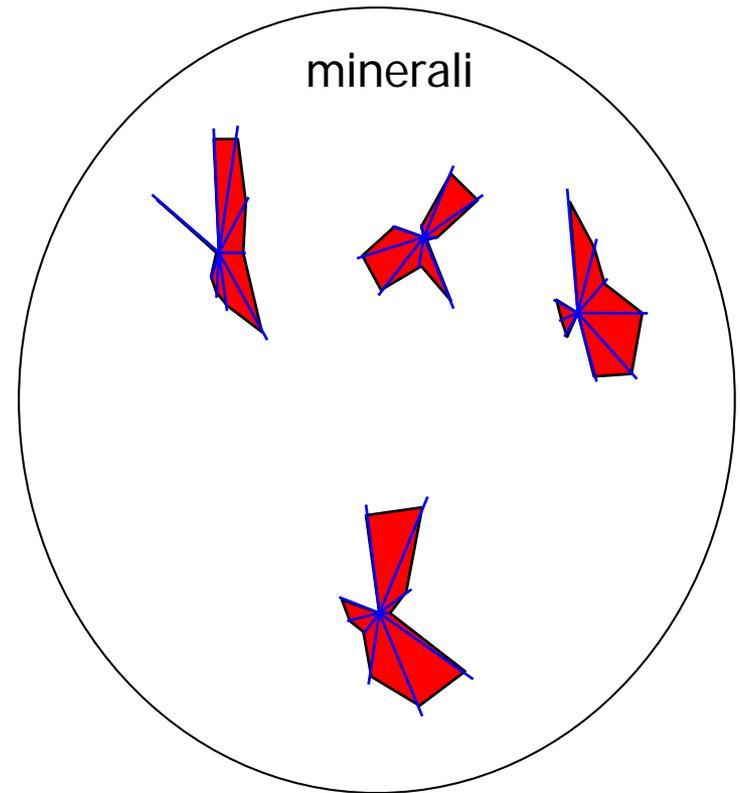
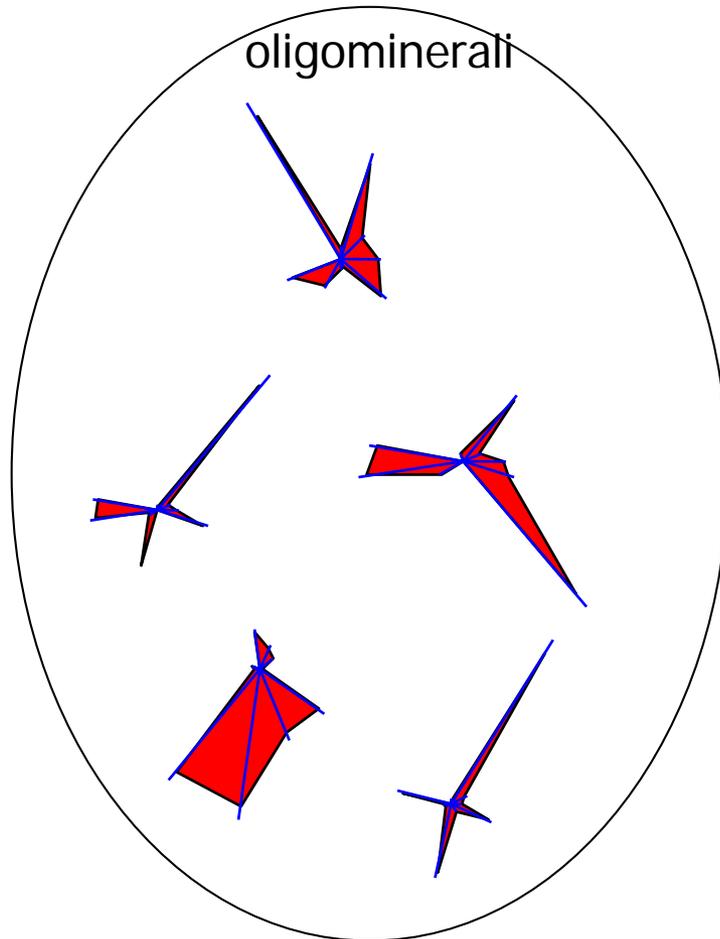
Radar plot

- Un modo semplice per visualizzare patterns multidimensionali
- Si definiscono N assi quanti sono le variabili del pattern
- Gli assi sono disposti a formare un cerchio
- Lungo ogni asse viene graficato il valore della variabile
- Unendo i punti sugli assi si ottiene una figura che forma il “profilo” del pattern
 - Utilizzato spesso in analisi sensoriale per definire i profili sensoriali degli alimenti.



;

Radar plot e class membership



Criteri di distanza

- Ogni pattern è un punto in uno spazio N-dimensionale
- Lo spazio è definito dalle features che descrivono il pattern
 - Ad ogni feature è assegnato un asse
 - Tutti gli assi definiscono una base orto-normale
- Relazione “class-membership” - distanza
 - Due punti (patterns) vicini appartengono verosimilmente alla stessa classe
 - due punti lontani appartengono a classi differenti

Criteri di classificazione

- Criteri “unsupervised” o esplorativi
 - Determinare, in base ad un criterio a-priori, uno schema di classificazione interno
 - Il criterio utilizzato è generalmente quello della distanza
- Analisi dei Clusters
 - Metodo di raggruppamento gerarchico dei campioni a formare classi via via meno definite.
- Metodi “esotici”
 - Metodo del potenziale

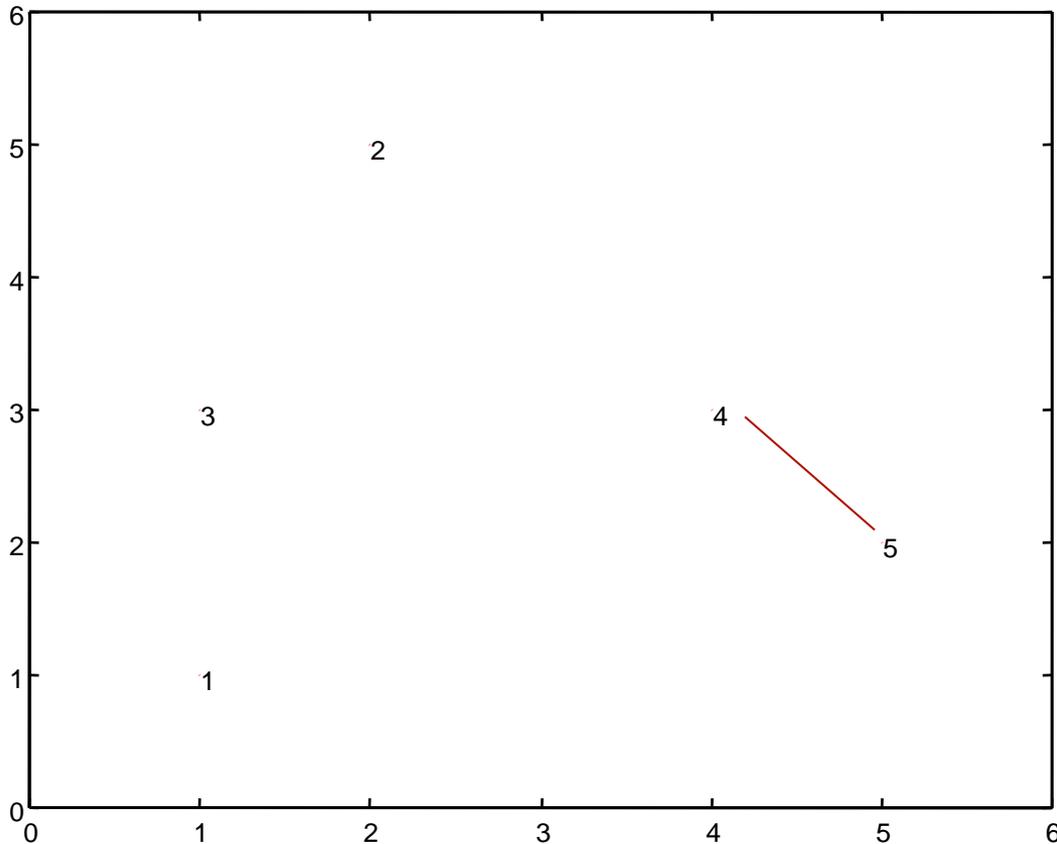
Cluster Analysis

- Dato un insieme di punti la cluster analysis forma una gerarchia di raggruppamento
- In funzione della distanza I punti si raggruppano a formare insieme via via più grandi.
- Alla fine tutti I dati formano un solo insieme
- Lo strumento base della cluster analysis è la matrice delle distanze
 - Dato un insieme di pattern X_i si definisce matrice delle distanze d_{ij} la seguente:

$$d_{ij} = \left\| X_i - X_j \right\|$$

- Dove l'operazione di norma di vettore identifica generalmente la norma euclidea
- La matrice è ovviamente simmetrica

Esempio: matrice di distanza



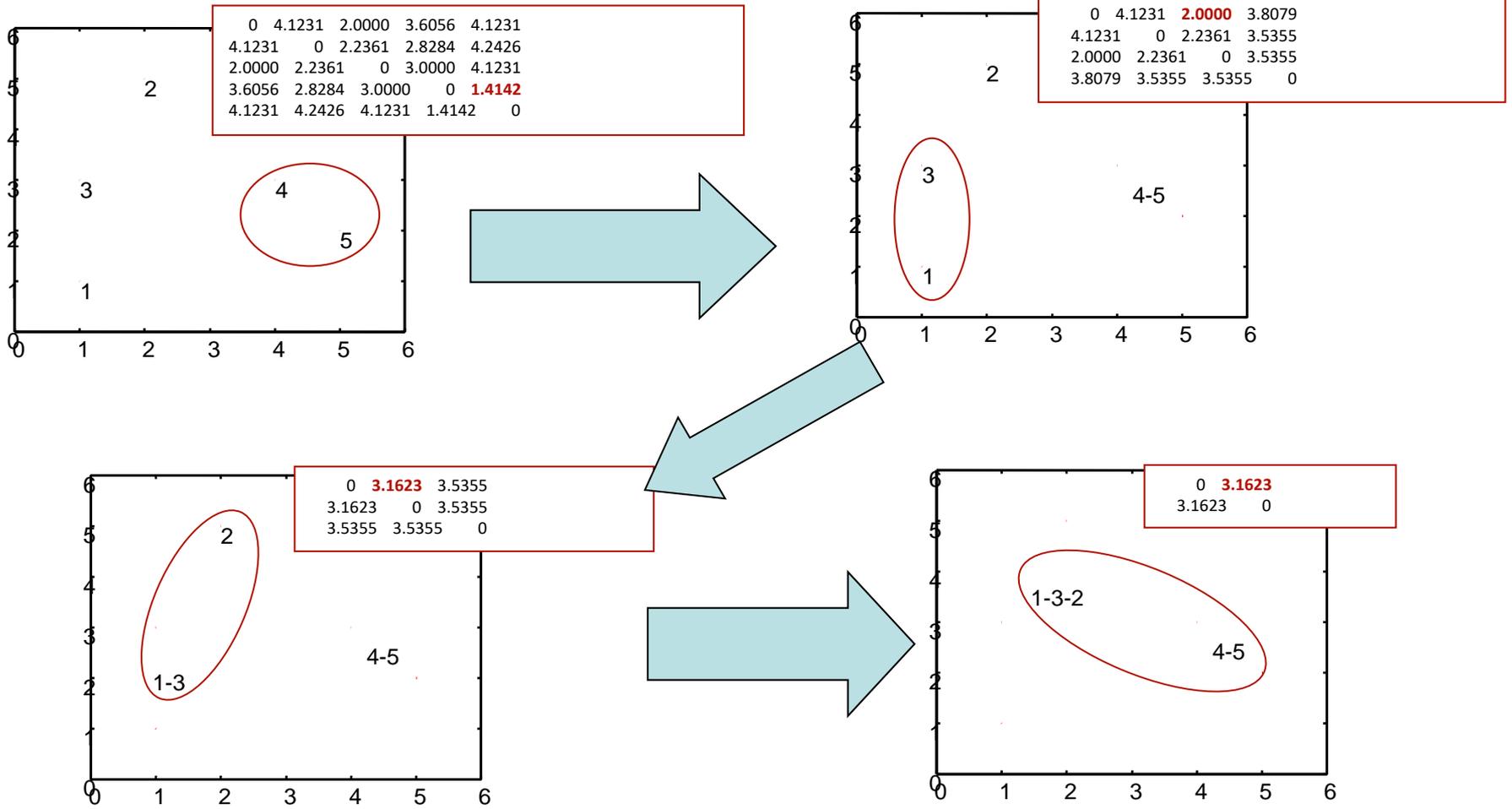
	1	2	3	4	5
1	0	4.1231	2.0000	3.6056	4.1231
2	4.1231	0	2.2361	2.8284	4.2426
3	2.0000	2.2361	0	3.0000	4.1231
4	3.6056	2.8284	3.0000	0	1.4142
5	4.1231	4.2426	4.1231	1.4142	0

- I punti 4 e 5 sono i più vicini
- La coppia 4-5 è isolata rispetto a 1-2-3
- Verosimilmente ci sono 2 gruppi di dati:
 - 1-2-3
 - 4-5
- La cluster analysis rende questi ragionamenti oggettivi e consente di operare su spazi N-dimensionali cioè con pattern di dimensioni qualsiasi

Cluster analysis

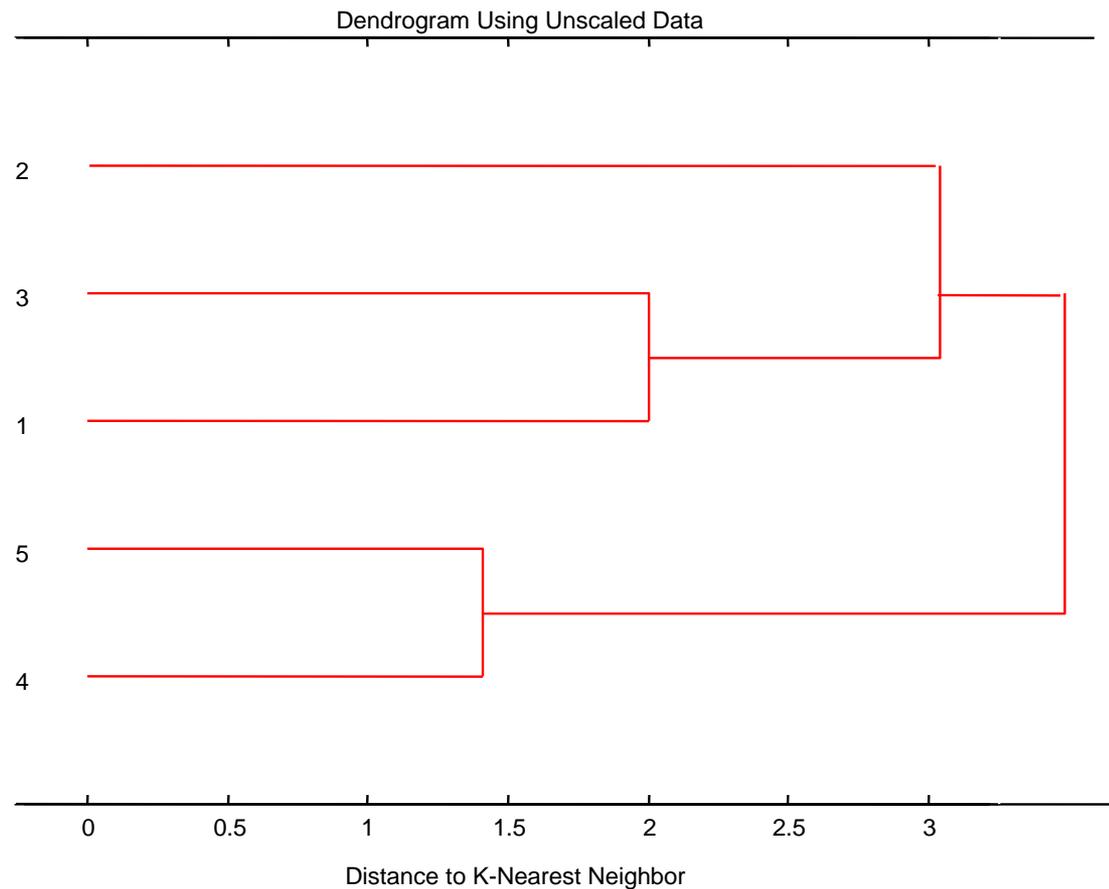
- La cluster analysis è un processo gerarchico iterativo che costruisce una struttura ad albero (dendrogramma) che definisce le classi in funzione della distanza.
 - Step i
 - Calcolo matrice di distanza
 - Individuazione della coppia di punti con distanza più piccola
 - Formazione del cluster unendo i punti
 - Sostituzione del cluster con il punto medio (o con uno dei punti della coppia)
 - Reiterazione del procedimento finchè rimane un solo punto

esempio

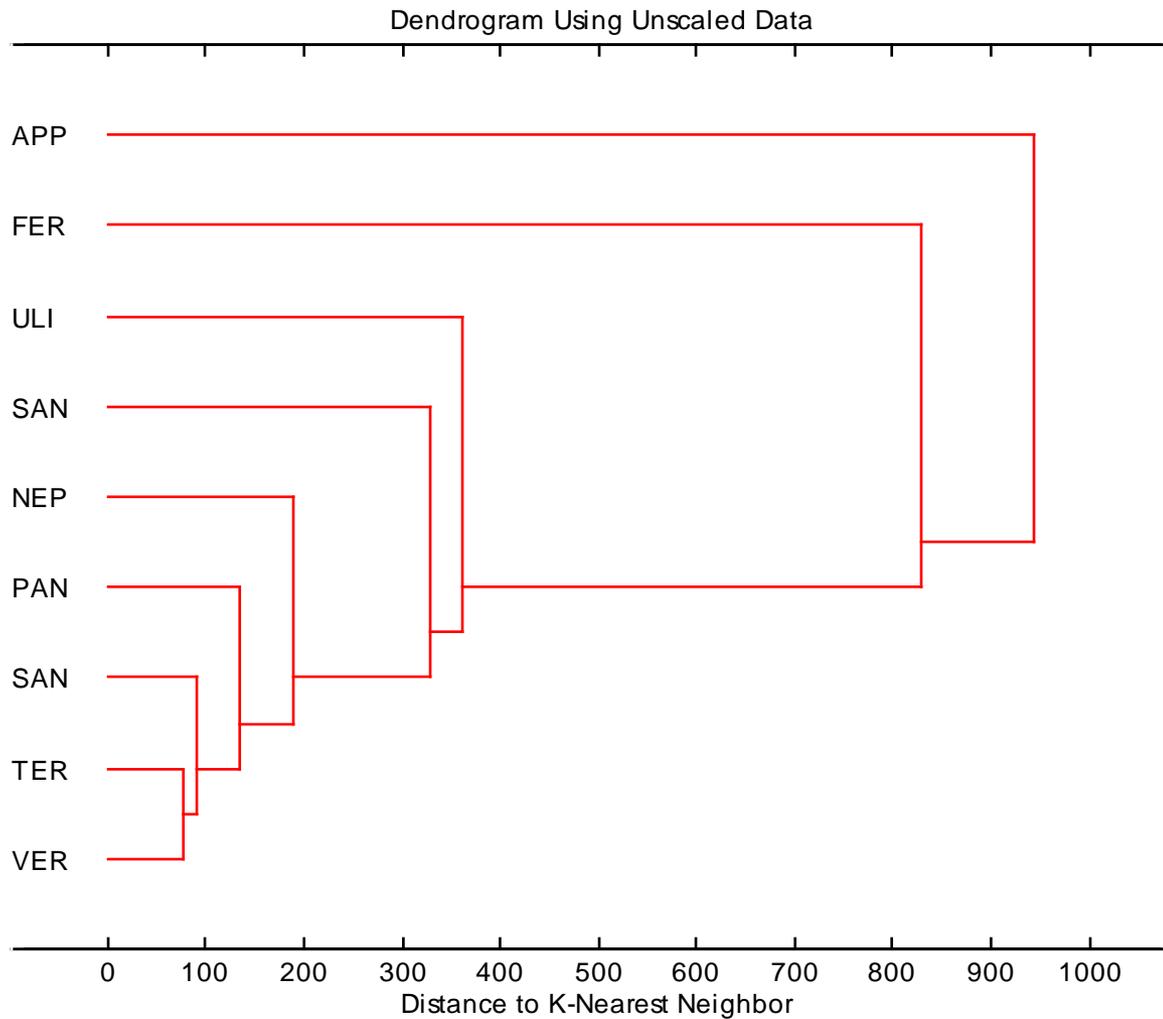


dendrogramma

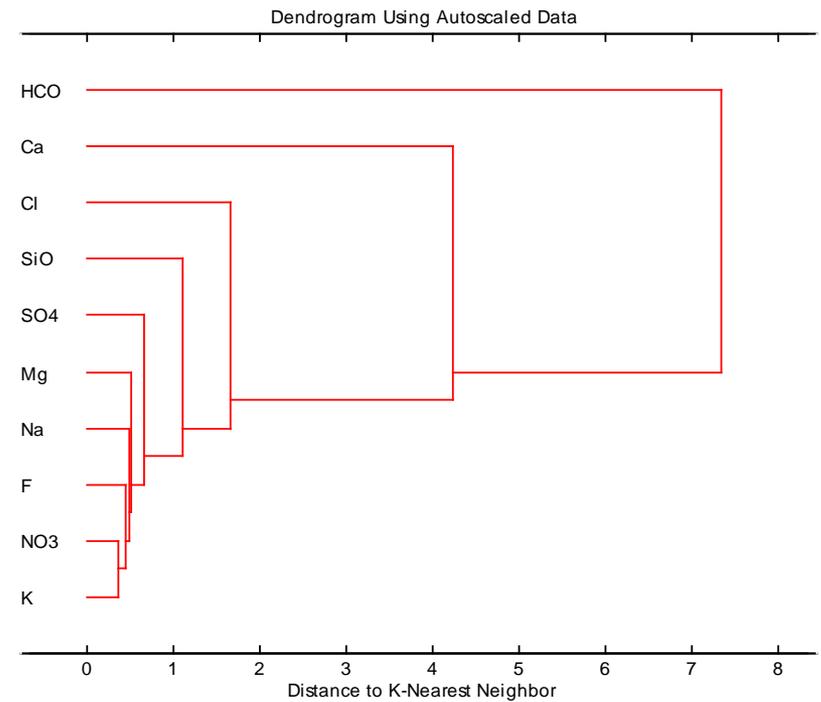
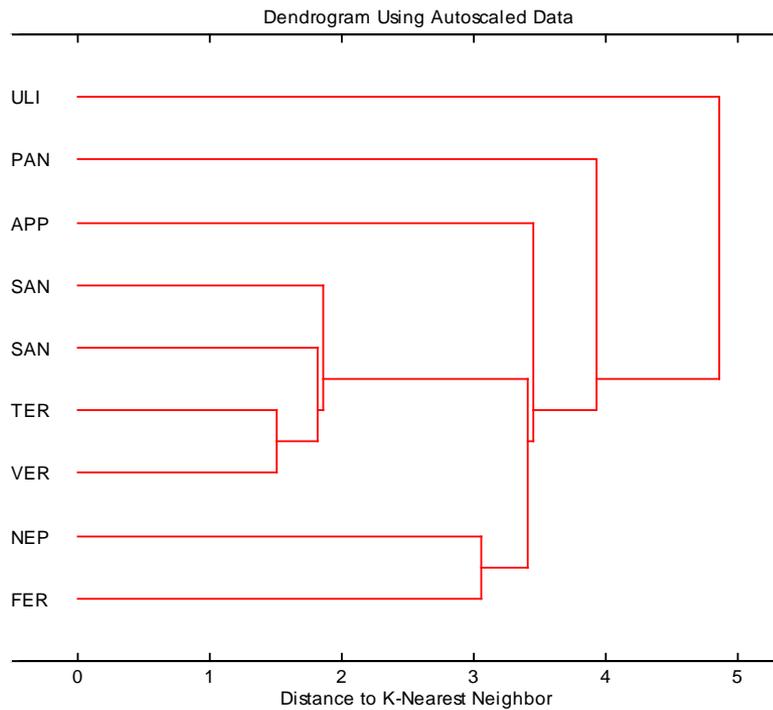
- Il risultato della cluster analysis viene raffigurato in un diagramma ad albero (dendrogramma) dove tutti i punti sono uniti, in funzione della distanza a formare dei cluster.



Cluster analysis esempio acque minerali

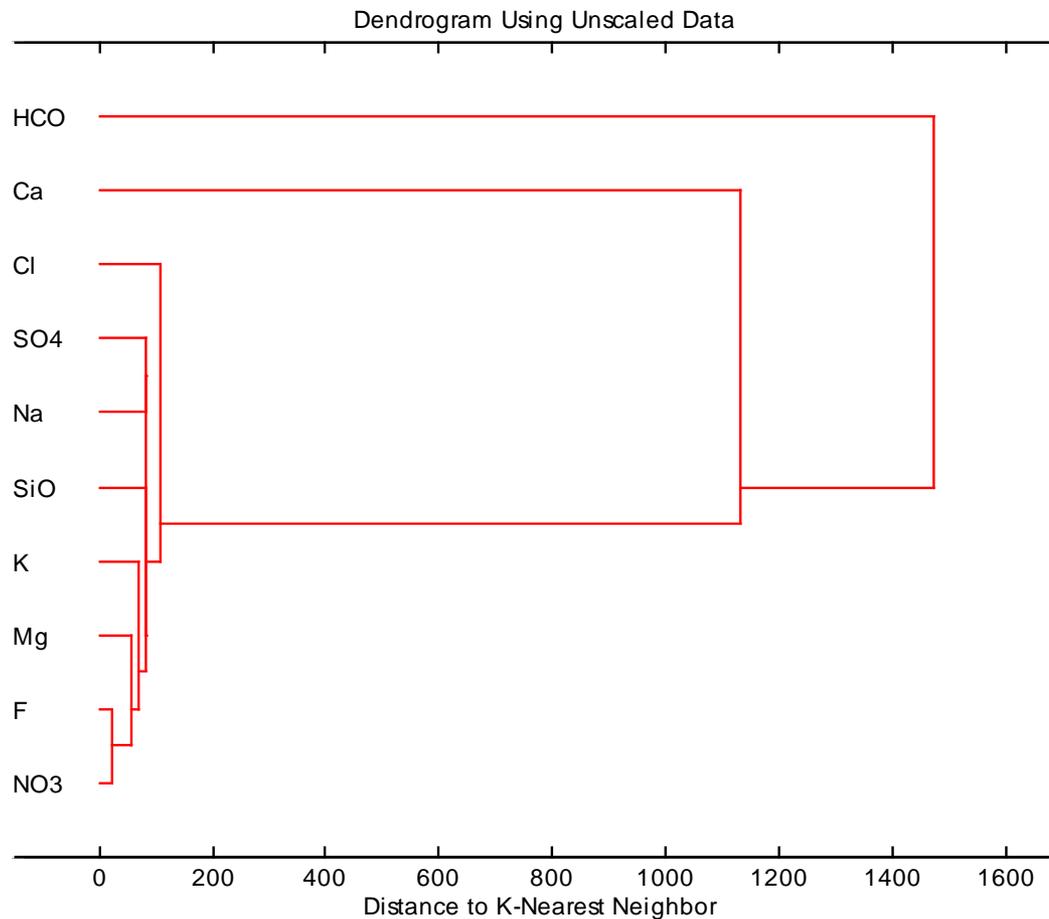


Cluster analysis con acque minerali normalizzate



Pattern recognition “inversa”

- Per studiare il ruolo delle features si può studiare il problema trasposto dove le features diventano i samples ed i samples le features.
- Esempio: cluster analysis con acque minerali

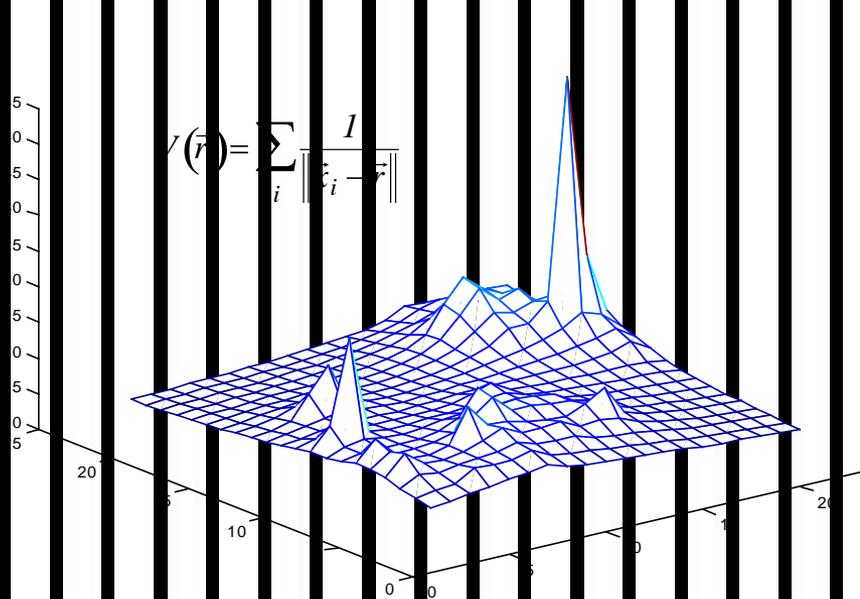
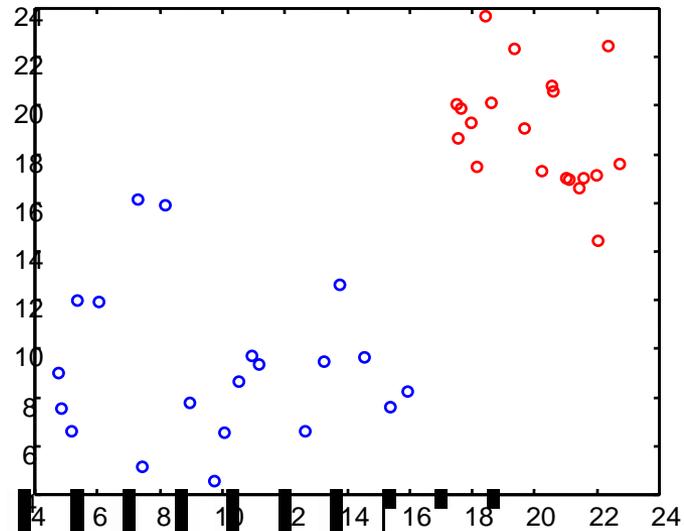


Metodo del Potenziale

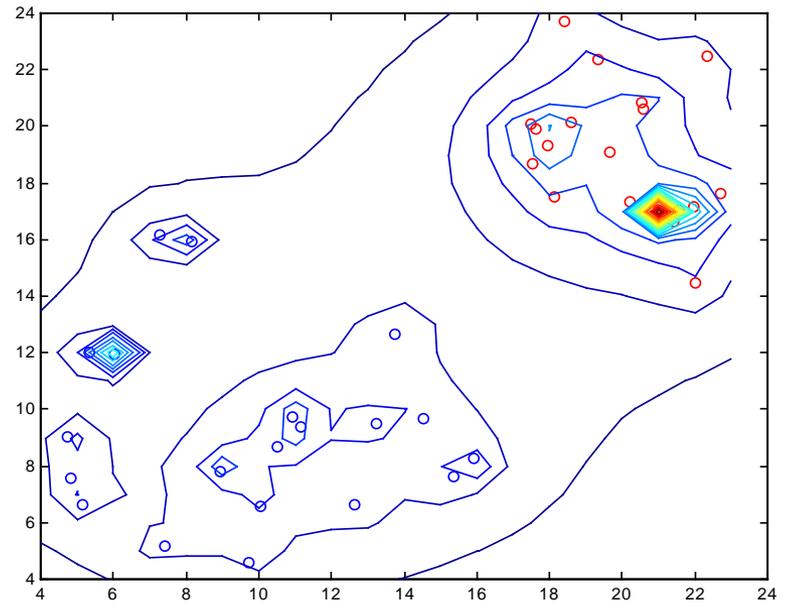
- Un metodo “esotico” interessante per la classificazione “unsupervised” è quello basato sulla analogia con I campi di forze.
- Supponiamo che ogni punto posseda una massa M (uguale per tutti). Tale massa genererà un potenziale V punto dello spazio. Dove I punti sono raggruppati (in una classe) si genererà un potenziale maggiore, studiando quindi l’andamento del potenziale, ed in particolare I suoi massimi è possibile individuare delle regioni spaziali di massimo addensamento (classi).
- L’analogia con le masse è solo un esempio si può utilizzare qualunque funzione potenziale abbia senso.
- Nel caso dell’analogia gravitazionale, dati N punti x_i il potenziale nel punto r è dato da:

$$V(\vec{r}) = \sum_i \frac{1}{\|\vec{x}_i - \vec{r}\|}$$

Esempio metodo del potenziale



$$V(\vec{r}) = \sum_i \frac{I}{\|\vec{x}_i - \vec{r}\|}$$



Partial Least Squares (PLS)

Partial Least Squares
PLS toolbox di MATLAB

Da PCR a PLS

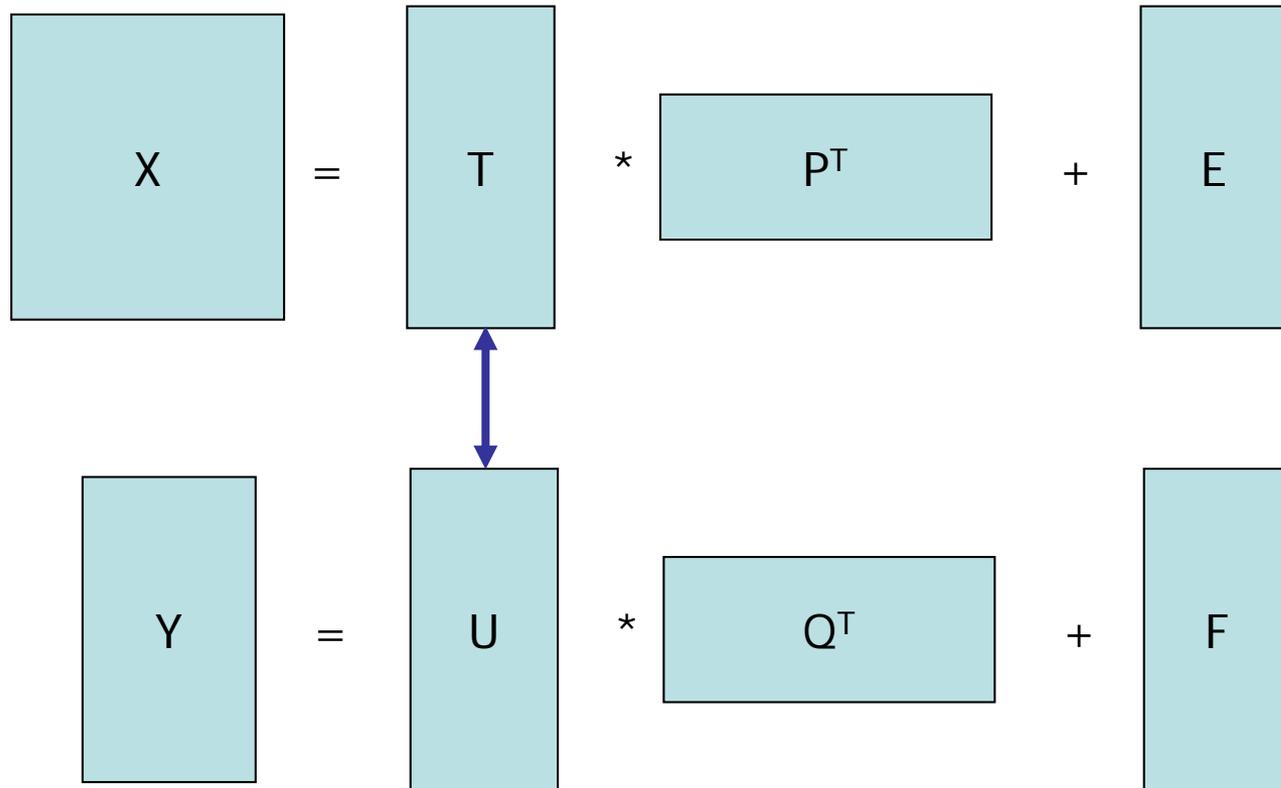
approccio geometrico

- Nella PCR la soluzione del problema della regressione passa attraverso la decomposizione della matrice dei dati nella matrice delle componenti principali
- Le componenti principali sono le direzioni, nello spazio delle variabili di X , che massimizzano la varianza e generano una base nella quale i dati di X risultano non correlati
- Nella PCR le componenti principali diventano le nuove variabili (non correlate) della regressione che così diventa *più facilmente* risolvibile.
- Nella PLS anche la matrice Y viene decomposta in componenti principali e le componenti principali di X vengono ruotate nella direzione di massima correlazione rispetto alle componenti principali di Y
- Scopo della PLS è determinare delle variabili latenti, simili alle componenti principali che massimizzano la varianza di entrambe le matrici

PLS e PCR

- PLS “funziona meglio” di PCR perchè le variabili latenti sono scelte massimizzando la correlazione con le PC della matrice Y
- Le componenti principali sono le direzioni di massima varianza, queste direzioni non dipendono dallo “scopo” (matrice Y) ma sono determinate univocamente dalla matrice X
- Le componenti principali quindi non sono di per se significative per la costruzione di un modello che stima Y, agevola solo la MLR eliminando la correlazione tra le variabili.

PLS calcolo variabili latenti



Le componenti principali sono determinate massimizzando la correlazione tra T e U e la loro varianza

$$\max [corr^2(U, T), var(U) \cdot var(T)]$$

Overfitting in PLS

- PLS è soggetta ad overfitting
- Il numero di variabili latenti deve essere ottimizzato in un processo di cross-validation
- L'overfitting è dato dal fatto data la variabile latente k la variabile $k+1$ è ottenuta fittando il sottospazio di dimensione $k+1$, poichè le variabili latenti non sono ortogonali tra loro, non c'è limite alla possibilità di fittare esattamente i dati di calibrazione con un numero N di variabili latenti.
- La cross-validation fissa il numero di variabili latenti la cui accuratezza è stimata sul set di validazione. Normalmente tale valore è più grande rispetto all'errore ottenuto dal modello sui dati di calibrazione
- Tali errori sono quantificati dalle grandezze
 - RMSEC Root Mean Square Error in Calibration
 - RMSECV Root Mean Square Error of Calibration in Validation

Algoritmo PLS

Modello PCA per X e Y

$$X = T \cdot P^T + E \quad Y = U \cdot Q^T + F$$

Determinazione delle variabili latenti

$$\max \left[\text{corr}^2(T, U), \text{var}(U) \cdot \text{var}(T) \right] \Rightarrow \begin{cases} T = X \cdot P = X \cdot X^T \cdot U \\ U = Y \cdot Q = Y \cdot Y^T \cdot T \end{cases}$$

Costruzione del modello: matrice di regressione B

$$Y = X \cdot B^T \Rightarrow B^T = \left(X^T X \right)^{-1} \cdot X^T \cdot Y$$

$$B^T = P \cdot \Lambda^{-1} \cdot T^T \cdot U \cdot Q^T$$

Matlab PLS toolbox *modlgui*

- Dati: 4 sensori TSMR, per la misura di ottano e toluene

Linear Regression

File Edit View Insert Tools Window Help

MODL_File

Var: s0,nt0
Data: modeled (calibration set)
Size: 24 by 4, 24 by 2
Samp Lbls:
Var Lbls:

Model: calibrated on loaded data
Method: SIMPLS
LV(s): 3
Data: 24 by 4, 24 by 2
Scaling: autoscaled

No. LVs: 3 show parameters

Latent Variable	Percent Variance X-Block		Percent Variance Y-Block	
	This LV	cum	This LV	cum
1	93.48	93.48	47.70	47.70
2	6.50	99.97	51.69	99.38
3	0.02	100.00	0.14	99.53
4	0.00	100.00	0.02	99.55

Buttons: calc, apply, plots, press, scores, loads, biplot, data

Regression Parameters

File Edit View Insert Tools Window Help

Scaling: autoscale

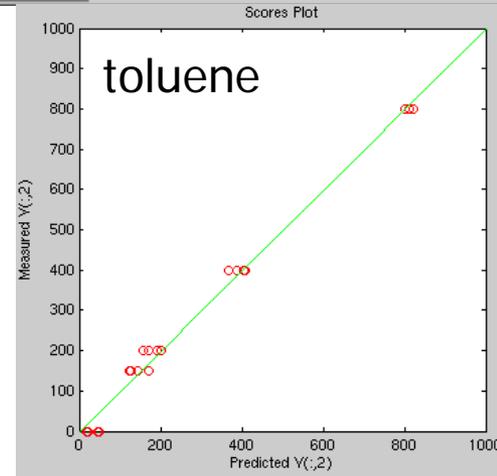
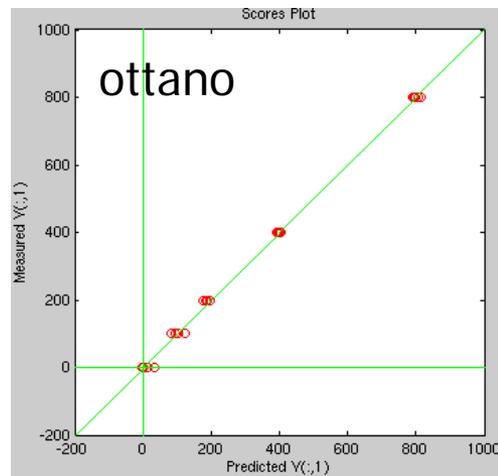
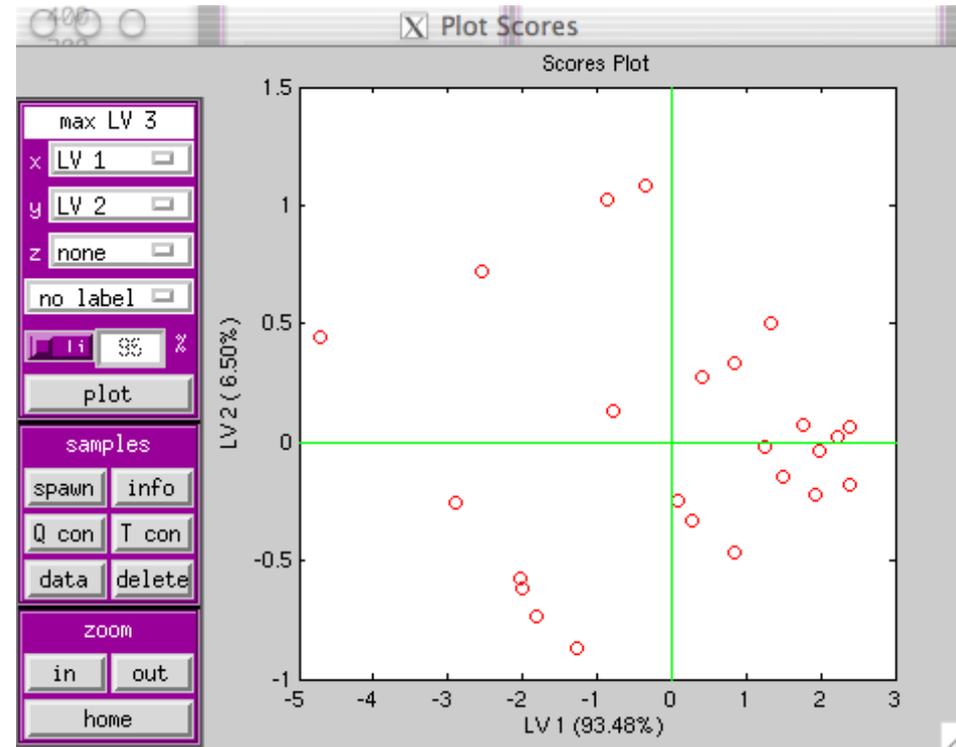
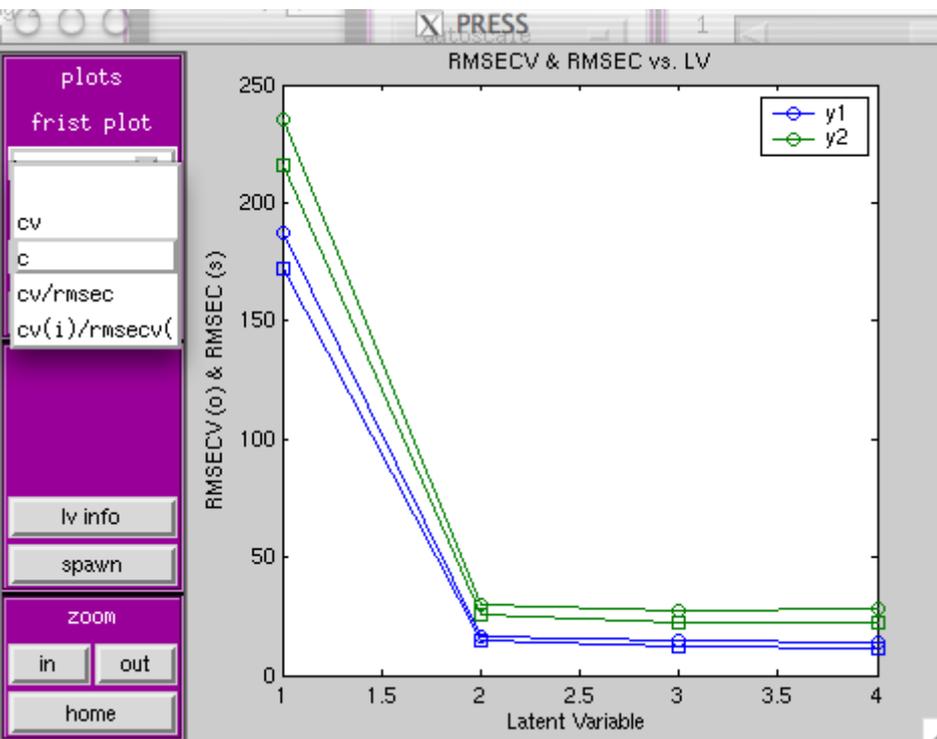
Regression: SIMPLS

Cross Validation: leave one out

Max LVs: 4

Cross Validation Options: leave one out, venetian blinds, contiguous block, random subsets

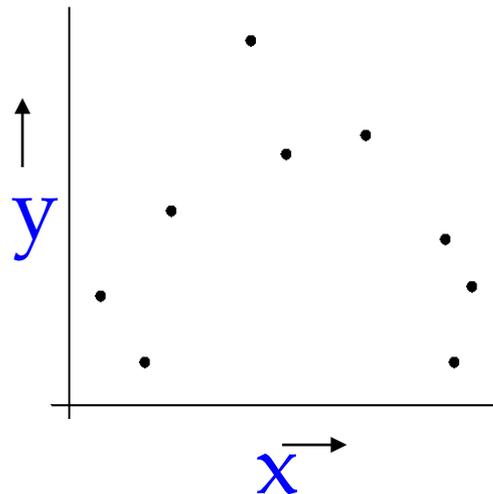
Esempio *modlgui*



Modello lineare o non-lineare?

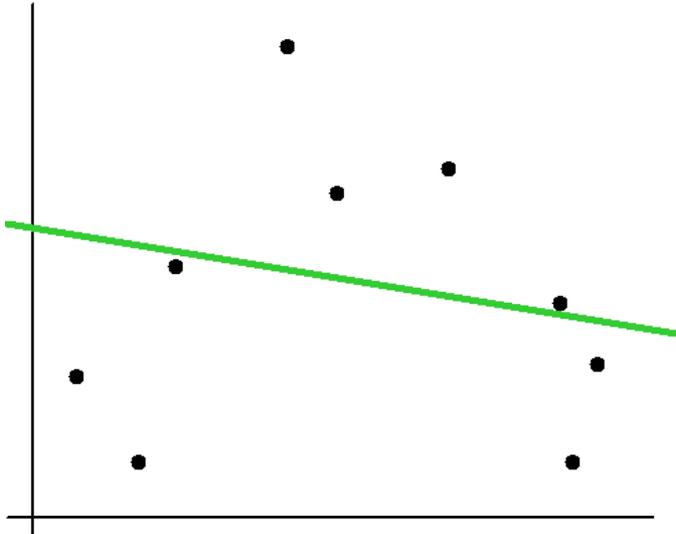
Il problema della validazione

- Quale è la migliore funzione che descrive i dati sperimentali?
- Quella che consente di predire con errore minimo le variabili che non sono state usate per costruire il modello.
- L'operazione che consente di stimare questo errore si chiama cross-validation.
- Esempio: consideriamo i seguenti dati in cui si ha: $y=f(x)+e$
 - Quale è la migliore funzione che descrive la relazione tra y e x ?

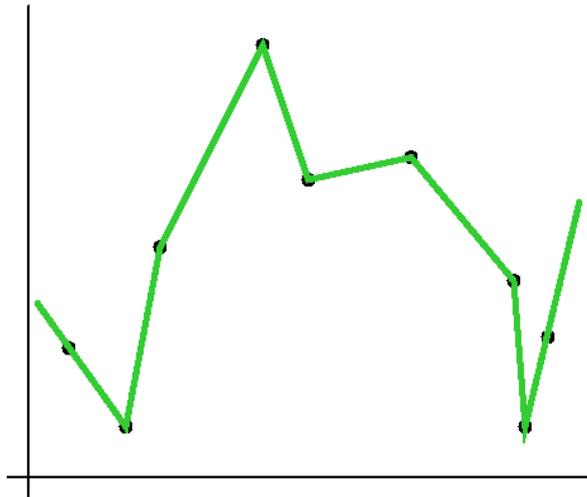
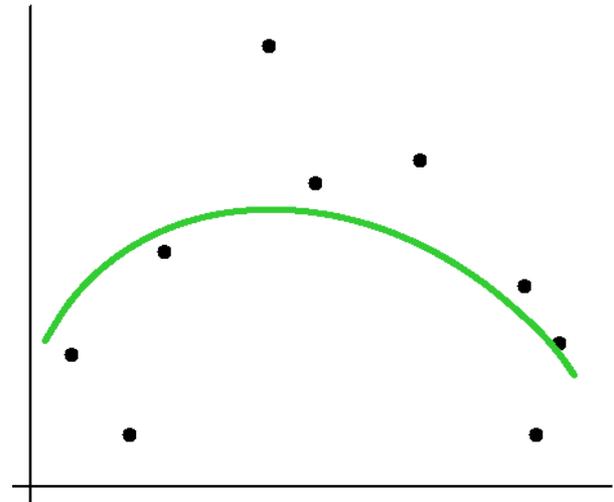


Esempi di soluzione

lineare



Moderatamente non lineare



Altamente non lineare

Il metodo del test

- L'insieme dei dati viene diviso in due
- Il modello viene determinato su un sottoinsieme dei dati (insieme di calibrazione *training set*)
- L'errore viene valutato sull'altro sottoinsieme (insieme di test *test set*)
- La stima dell'errore di predizione dell'insieme di test è significativa della capacità di generalizzazione del modello. Cioè della sua applicabilità a dati non utilizzati per la calibrazione. Quindi all'utilizzo del modello nel mondo reale per stimare grandezze incognite.

Stimatori della prestazione della regressione

- PRESS- Predicted Sum of Squares

$$PRESS = \sum_i (y_i^{LS} - y_i)^2$$

- RMSEC - Root Mean Square error of calibration

$$RMSEC = \sqrt{\frac{PRESS}{N}}$$

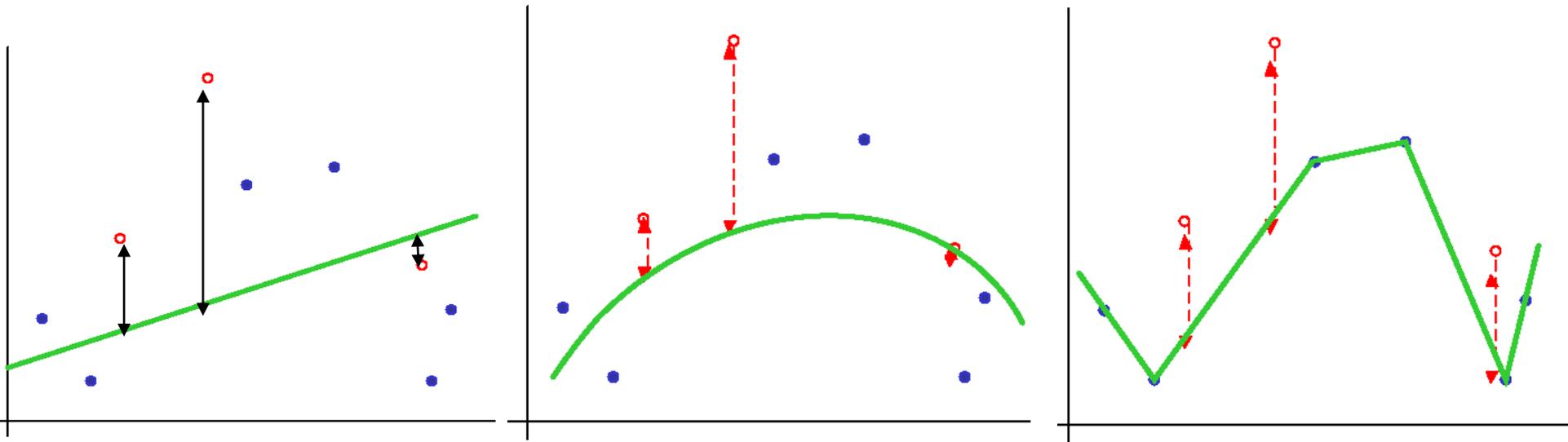
- RMSECV - Root Mean Square error of Cross-Validation
 - Per un modello che include k campioni non inclusi nel modello stesso

$$RMSECV_k = \sqrt{\frac{PRESS_k}{N}}$$

Applicazione del metodo del test

I dati segnati in rosso sono il test set. Il modello è calcolato sui restanti dati (punti blu).

L'errore sui dati di test è valutato come RMSECV



RMSECV=2.4

RMSECV=0.9

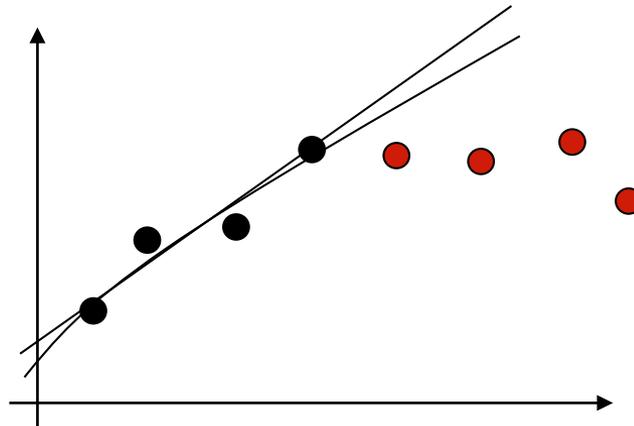
RMSECV=2.2

Discussione

- Il metodo migliore è quello moderatamente non lineare (quadratico)
- Il metodo lineare ha un errore grande sia in calibrazione che in test
- Il metodo fortemente non lineare ha un errore di calibrazione nullo ma un errore di test elevato. Tale modello è “troppo specializzato” nel descrivere i dati di calibrazione e non è in grado di generalizzare cioè di trattare adeguatamente i dati non usati per la calibrazione.
 - Tale effetto si chiama **overfitting** ed è tipico nel caso di modelli fortemente non lineari.
 - Inciso: dati n punti questi sono fittati perfettamente da un polinomio di ordine n

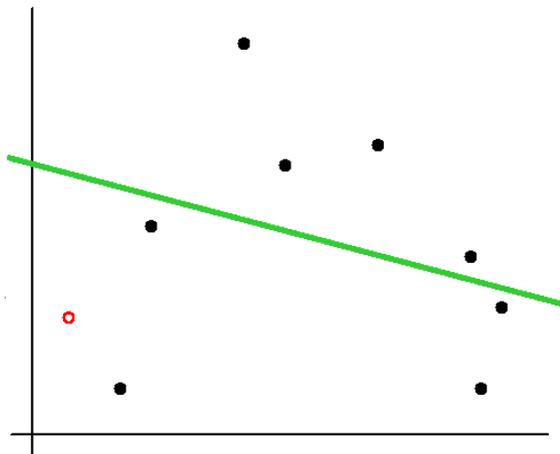
Considerazioni sul metodo del test-set

- Il metodo è molto semplice ma richiede insiemi di dati numerosi.
- La selezione degli insiemi non è semplice in generale va fatta in maniera casuale ma bisogna evitare di sbilanciare i due insiemi
 - È bene controllare che i due insiemi abbiano la stessa varianza e la stessa media
 - Se i due insiemi sono disgiunti si possono verificare fenomeni di overfitting apparente
 - Inciso: a parte casi semplici, in genere i modelli falliscono nelle estensioni analitiche cioè nella predizione di misure fuori del range considerato per la calibrazione.



Leave-One-Out cross-validation

- Nel caso di insiemi numerosi di dati il metodo del test set si può utilizzare senza problemi, ma quando il numero di dati diventa esiguo è necessario ricorrere ad altre strategie per la scelta della funzione e per la stima dell'errore.
- Il metodo più utilizzato in tal senso è il leave-one-out
 - Letteralmente *lasciane uno fuori*
- Il metodo consiste nel ridurre a 1 il contenuto dell'insieme di test, e nel valutare l'errore di predizione rispetto al dato eliminato. Eliminando consecutivamente tutti i dati e facendo la media degli errori ottenuti si ottiene una stima robusta dell'errore di predizione del modello funzionale.



Passo i-esimo:

Escludere il dato i

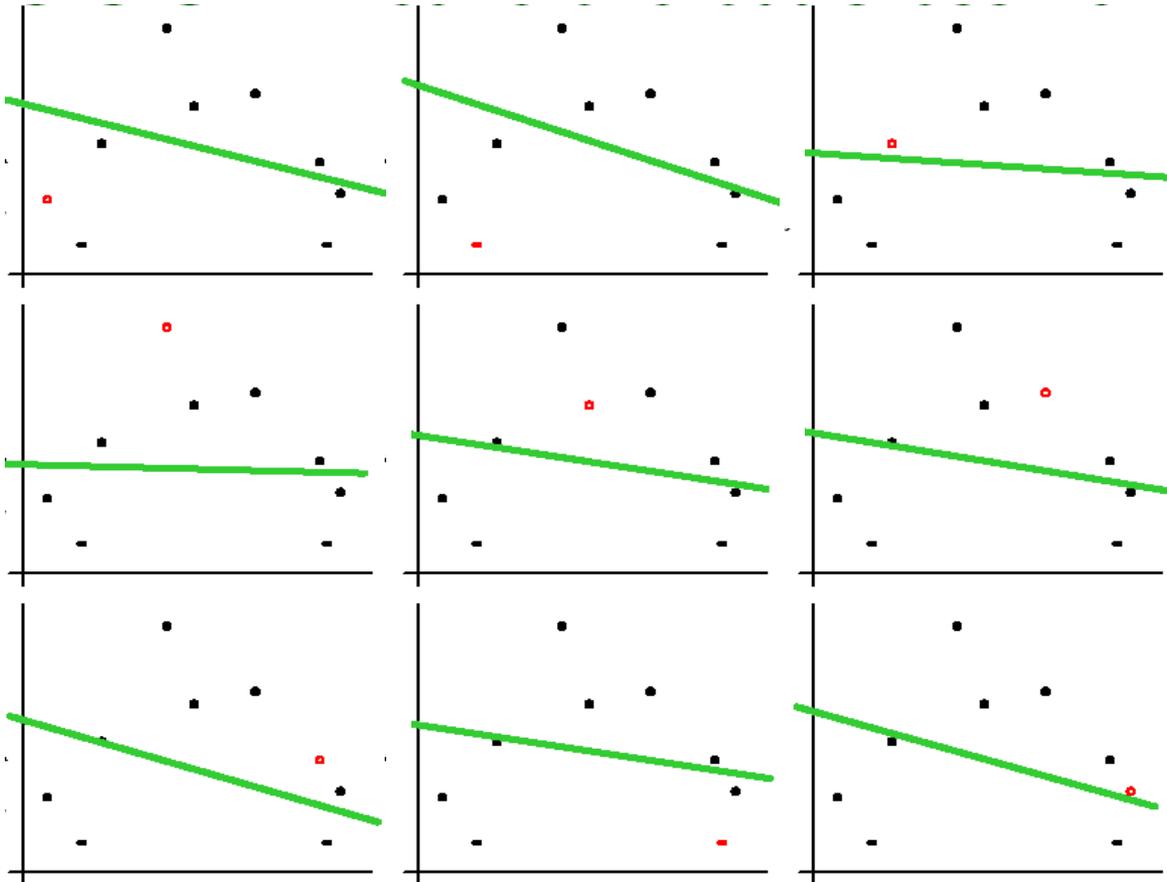
Calcolare la regressione sugli $n-1$ dati

Valutare l'errore di predizione rispetto al dato i -esimo

Immagazzinare il valore in ϵ_i

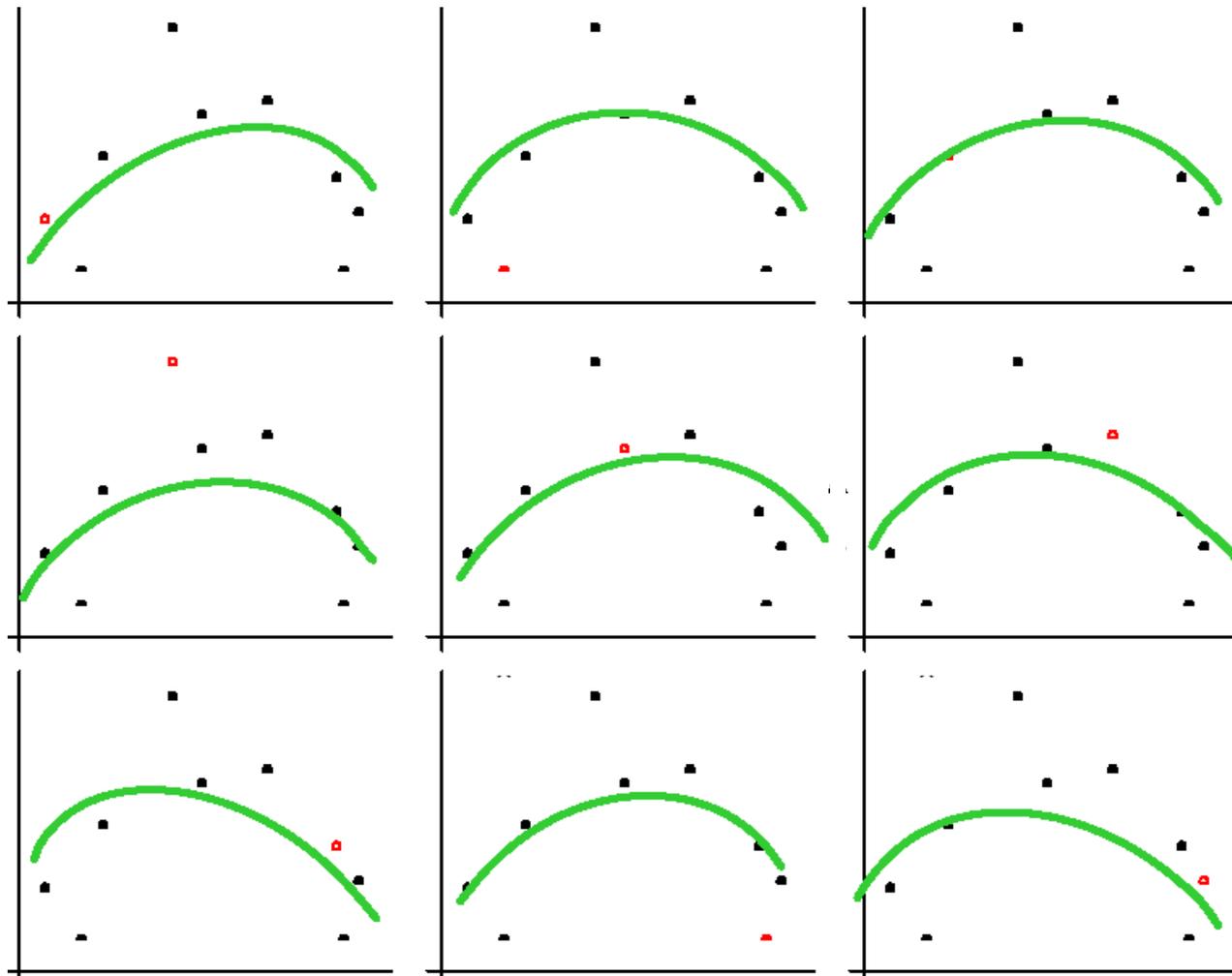
L'errore finale di predizione è la media degli ϵ_i

LOO modello lineare



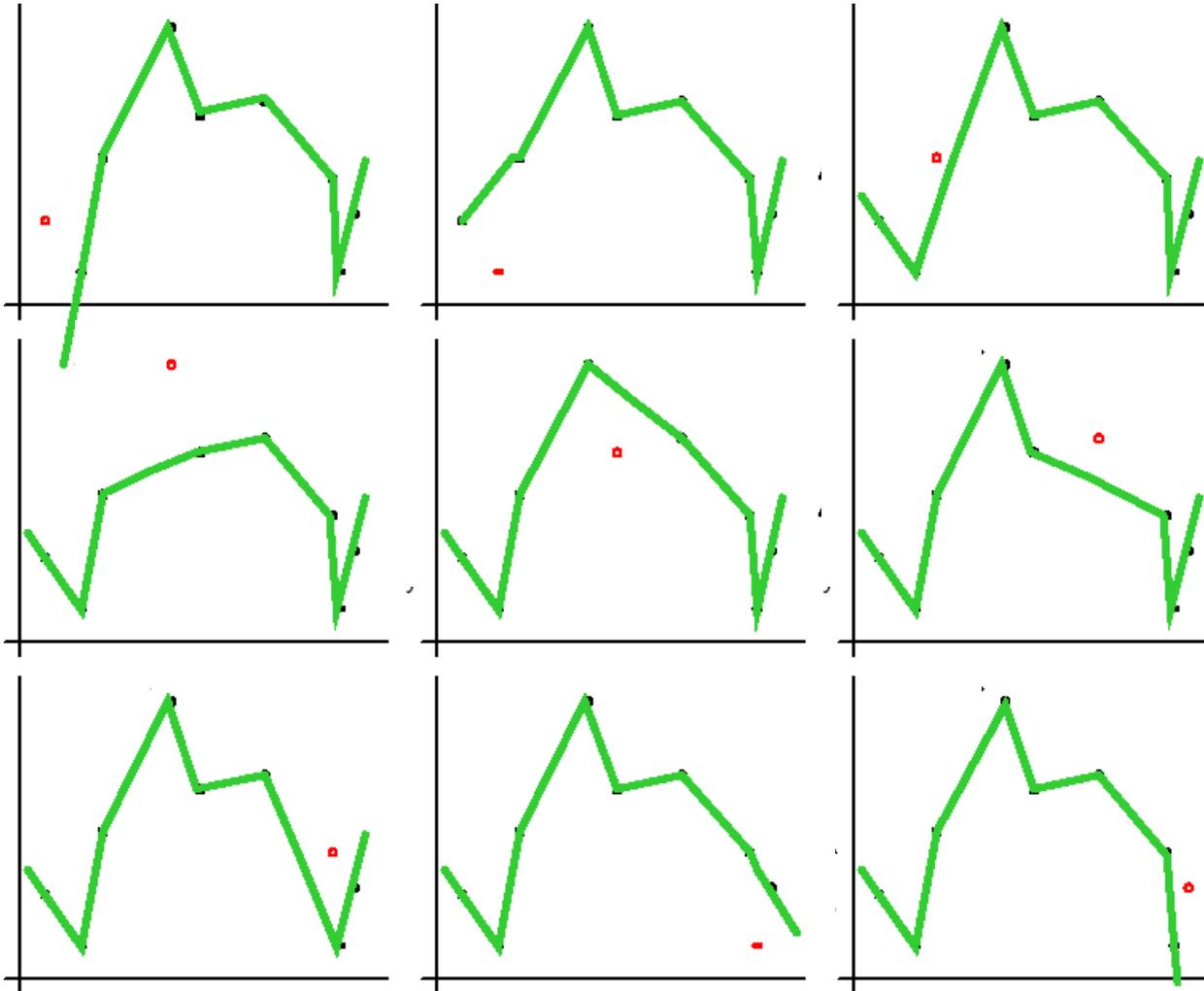
RMSECV=2.12

LOO modello moderatamente non lineare (quadratico)



RMSECV=0.96

LOO modello altamente non lineare



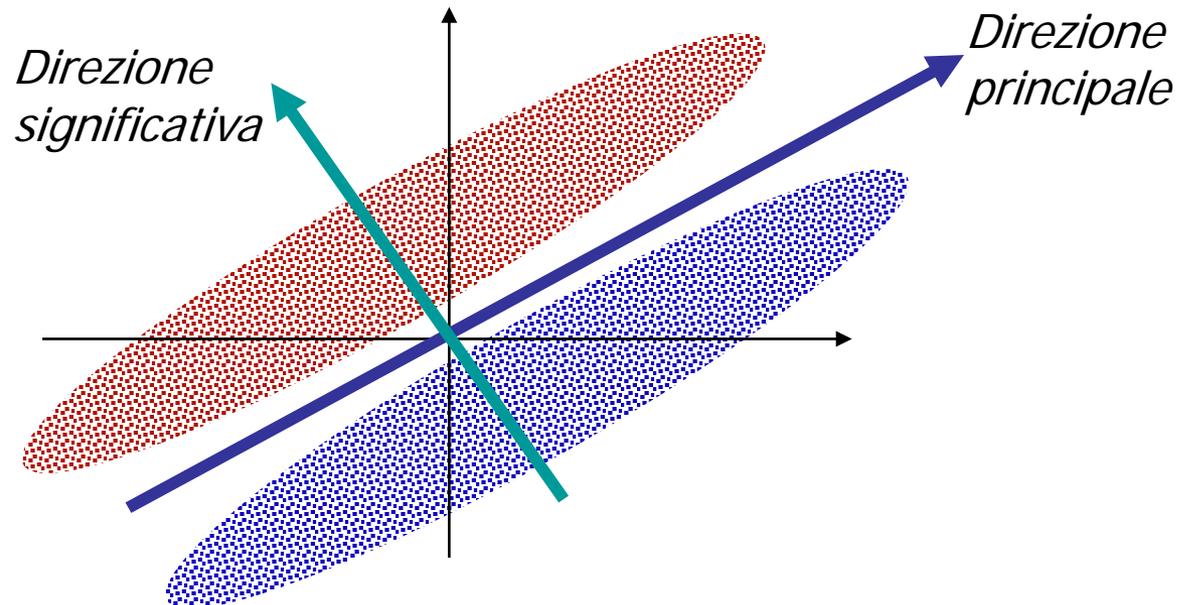
RMSECV=3.33

Confronto test - LOO

- LOO fornisce una stima migliore dell'errore di predizione rispetto al test set la cui stima dell'errore è inattendibile.
- LOO sfrutta al massimo l'intero set di dati.
- Ovviamente LOO è il metodo del test set con insieme di validazione di ampiezza minima.
- Per insiemi di grandi dimensioni LOO è dispendioso dal punto di vista del calcolo.
- Può essere "ammorbidito" considerando più insiemi di k dati.

Componenti principale e direzioni "significative"

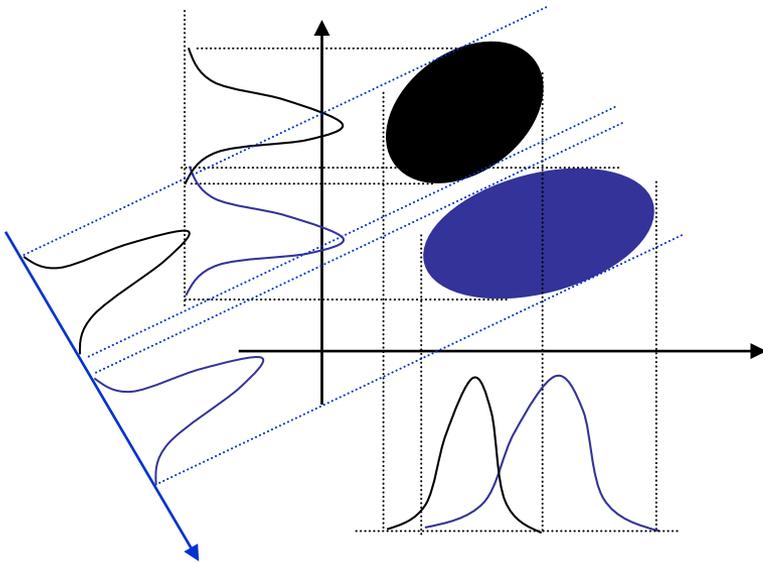
- Le componenti principali sono gli assi principali dell'ellissoide relativo alla matrice di covarianza, nulla assicura il fatto che queste direzioni siano importanti per il problema in esame
- La direzione "importante" definisce un punto di vista "supervised" cioè orientato a mettere in evidenza alcune proprietà del set di patterns



Analisi Discriminante Lineare (LDA)

- Un metodo che cerca uno o più direzioni lungo le quali viene massimizzata la separazione tra classi.

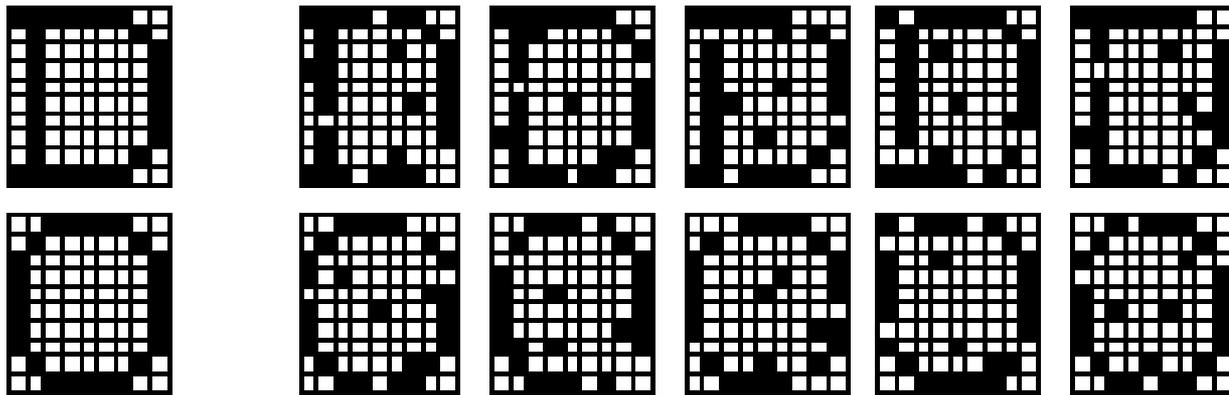
Significato Geometrico



- Esiste una classe di vettori di base (diversi dalle PC) dove la separazione tra classi è massima
- Se ci sono più classi si possono introdurre più direzioni
- Le direzioni discriminanti sono combinazione lineare delle variabili reali, si può studiare il contributo di ogni variabile alla direzione discriminante.

Template Matching

- Questo metodo è efficiente quando ogni classe contiene solo un pattern. I pattern misurati sono affetti da rumore additivo (no traslazione, rotazione, deformazione del pattern)
 - Esempio tipico: Optical Character Recognition
- Per ogni classe, il pattern con minore rumore è considerato come template della classe
- Due modalità di assegnare un pattern ad una classe:
 - Contare il numero di accordi: correlazione massima
 - Contare il numero di disaccordi: errore minimo



templates

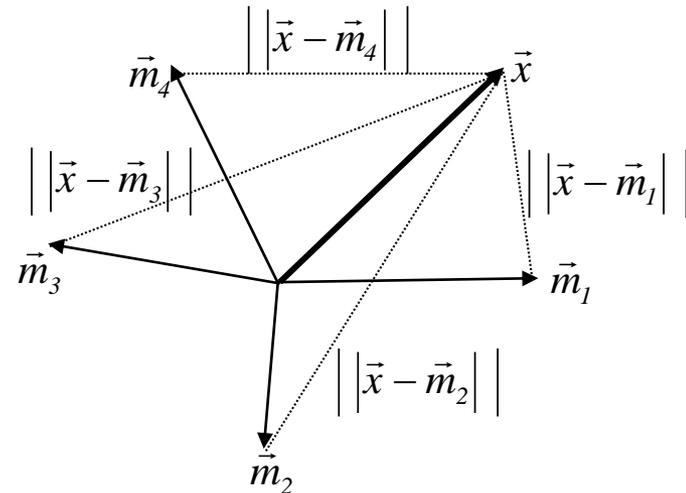
Measured Patterns

Classificatore di distanza minima (K-nearest neighbour)

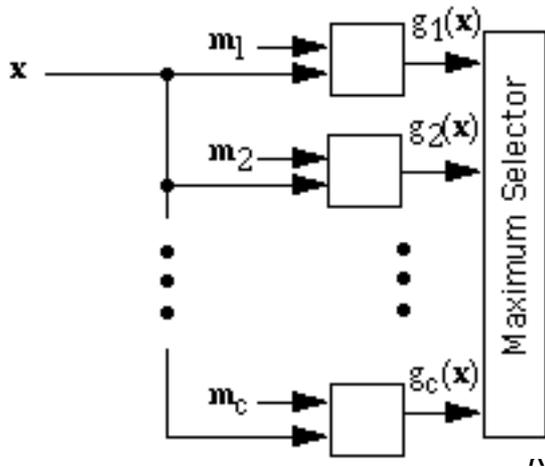
- Espressione matematica del template matching
- I templates sono rappresentati da vettori \mathbf{m}_j
- Per la classe K, la distanza tra un pattern(\mathbf{x}) ed il template corrispondente (\mathbf{m}) è:

$$\mathcal{E}_k = \left\| \vec{\mathbf{x}} - \vec{\mathbf{m}}_k \right\|$$

- Il pattern \mathbf{x} è assegnato alla classe verso la quale la distanza dal template corrispondente è minima
- L'operazione richiede la definizione di una metrica
 - Cioè di una regola per il calcolo delle distanze



funzioni discriminanti



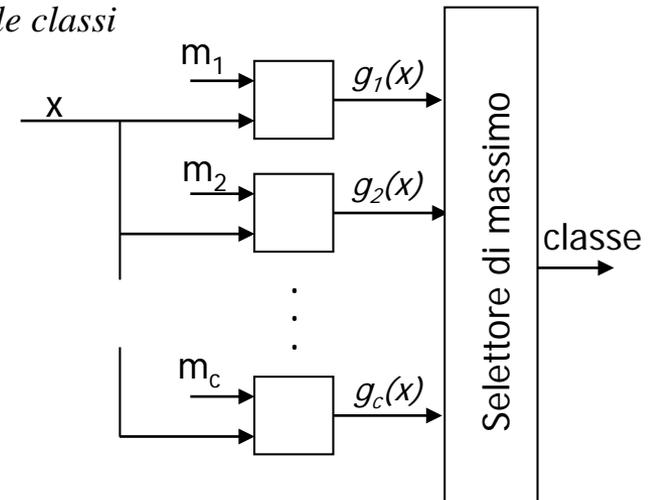
Classifier:

$$\vec{n}_k)^T (\vec{x} - \vec{m}_k)$$

$$\vec{n}_k + \vec{m}_k^T \vec{m}_k =$$

$$= -2 \left[\vec{m}_k^T \vec{x} - \frac{1}{2} \vec{m}_k^T \vec{m}_k \right] + \vec{x}^T \vec{x}$$

Uguale per tutte le classi



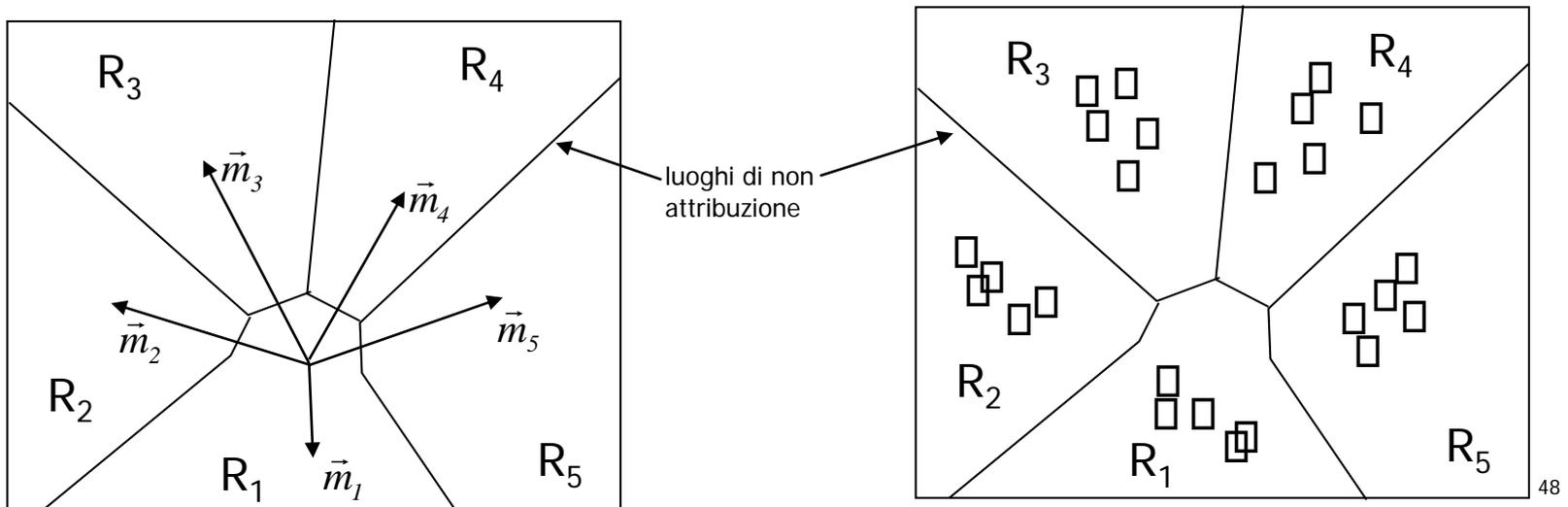
- Minimizzare l'errore significa rendere massima la seguente:

$$g(\vec{x}) = \left[\vec{m}_k^T \vec{x} - \frac{1}{2} \vec{m}_k^T \vec{m}_k \right]$$

- Questo metodo può essere interpretato come un classificatore di massima correlazione
- Le funzioni $g(x)$ sono dette funzioni discriminanti.

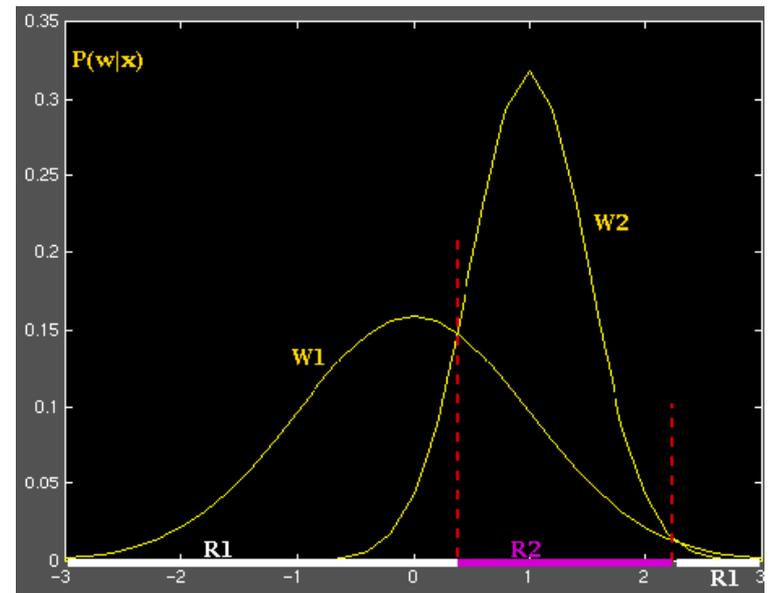
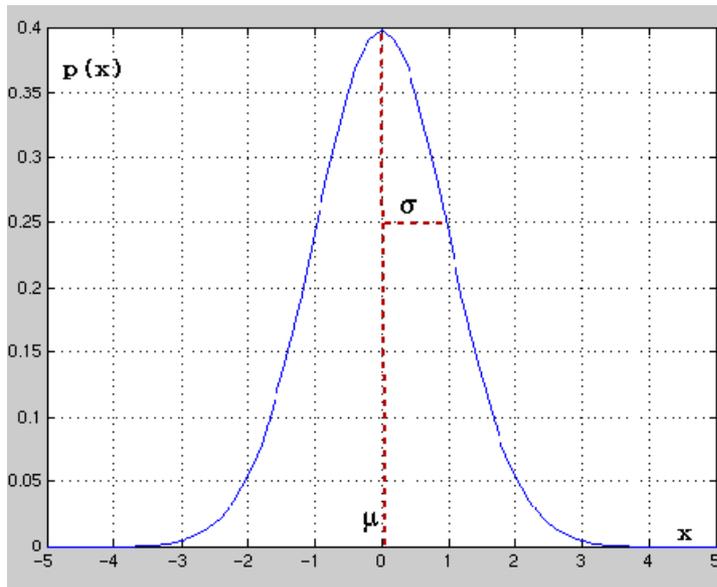
Contorni di decisione

- Le k funzioni discriminanti partizionano lo spazio in regioni ciascuna attinente ad una classe
- I contorni sono I luoghi di punti che giacciono alla stessa distanza da due o più template
 - Spazi di non decisione
- Questa operazione è detta tessellazione. Funzioni discriminanti lineari producono contorni poligonali.



Classificatore statistico

- Abbiamo visto nei casi precedenti come ogni classe sia definita da un solo pattern che funge da template per l'identificazione
- In pratica, classi complesse sono definite da una molteplicità di patterns
- Supponiamo di essere in grado di stimare, dai dati sperimentali, la distribuzione di probabilità dei patterns all'interno di ciascuna classe.
- Le funzioni di probabilità permettono di calcolare i contorni di decisione come le curve di iso-probabilità lungo le quali cioè la probabilità di appartenere a due o più classi sono uguali,
- Il caso più semplice è quello della distribuzione normale
 - Esempio: una dimensione (patterns formati da una sola feature):

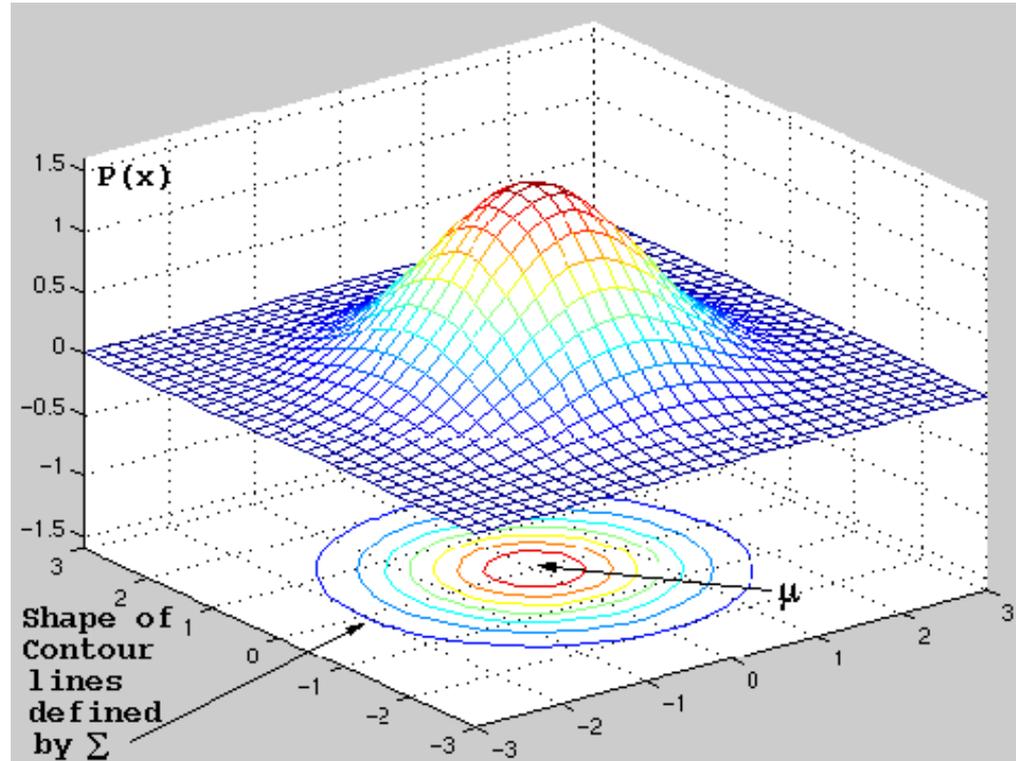


Distribuzione Multi-Normale

- $\boldsymbol{\mu}$ è il vettore di media
- La varianza (σ) è sostituita dalla matrice di covarianza (Σ).
- Il luogo dei punti di isoprabilità è la quadrica generata da Σ

$$\boldsymbol{\mu} = \begin{bmatrix} \mu_1 \\ \vdots \\ \mu_d \end{bmatrix} \quad \Sigma = \begin{bmatrix} \sigma_{11} & \cdots & \sigma_{1d} \\ \vdots & & \vdots \\ \sigma_{d1} & \cdots & \sigma_{dd} \end{bmatrix}$$

$$p(\mathbf{x}) = \frac{1}{\sqrt{2\pi} \sqrt{\Sigma}} \exp \left[-\frac{1}{2} (\mathbf{x} - \boldsymbol{\mu})^t \Sigma^{-1} (\mathbf{x} - \boldsymbol{\mu}) \right]$$

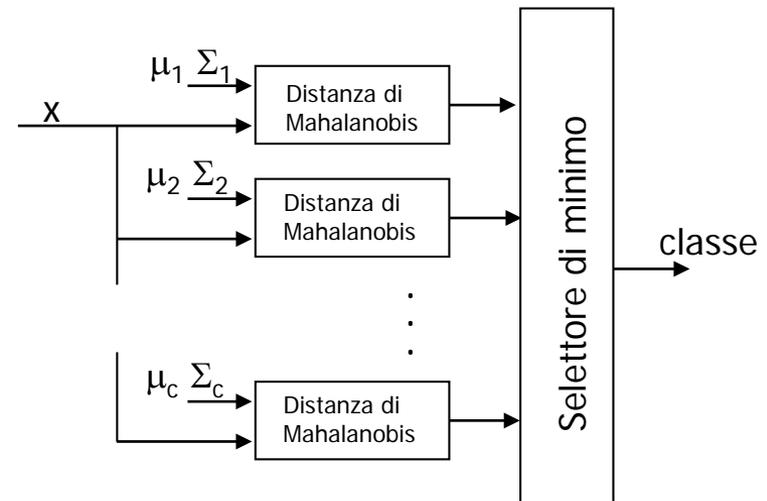


Mahalanobis distance

- Data una distribuzione gaussiana, il luogo di punti di iso-probabilità è dato dalla seguente forma quadratica:

$$(\mathbf{x} - \boldsymbol{\mu})^T \boldsymbol{\Sigma}^{-1} (\mathbf{x} - \boldsymbol{\mu})$$

- Il valore della probabilità è detto distanza di Mahalanobis, o distanza statistica, definisce la probabilità che ha un vettore \mathbf{x} di appartenere ad una distribuzione gaussiana definita da $\boldsymbol{\mu}$ e $\boldsymbol{\Sigma}$
- Valutando la distanza di Mahalanobis di un pattern verso ogni classe di statistica definita è possibile assegnare il pattern alla classe verso la quale la distanza di Mahalanobis è minore, cioè verso la quale è maggiore la probabilità di appartenenza.



Classificatori non parametrici

- In generale la distribuzione di probabilità dei patterns di ogni classe non è nota oppure non è gaussiana.
- In questo caso conviene utilizzare un classificatore non-parametrico cioè indipendente dalla statistica delle classi e nelle quali si cerca una combinazione delle variabili che consente l'identificazione delle classi.
- Questa operazione è simile alla analisi discriminante
- Il calcolo è simile alla multiple linear regression, nella quale vengono messe in relazione lineare la matrice X (insieme dei patterns dei campioni misurati) e una matrice Y (codifica numerica della class membership)
- Il problema consiste quindi nel trovare la matrice di regressione B che risolve il problema

$$Y = X \cdot B^T \Rightarrow B^T = \left(X^T \cdot X \right)^{-1} \cdot X^T \cdot Y$$

Codifica "one-of-many"

- La matrice Y è costruita con un numero di colonne pari al numero delle classi nel problema
- Ogni colonna di Y quindi identifica una classe
- Dato un pattern X_i la corrispondente riga Y_i è costruita ponendo a zero tutti gli elementi tranne quello corrispondente alla classe di X_i che viene posto uguale a 1
- In fase di identificazione il modello di regressione avrà una accuratezza finita, quindi il pattern viene assegnato alla classe il cui valore corrispondente risulta più grande.

PLS-Discriminant Analysis (PLS-DA)

- PLS è lo strumento ideale per la soluzione dei problemi di classificazione lineare.
- Oltre a minimizzare l'errore di classificazione, attraverso lo score ed il loading plot è possibile studiare quali sono le variabili del pattern che contribuiscono maggiormente alla classificazione.

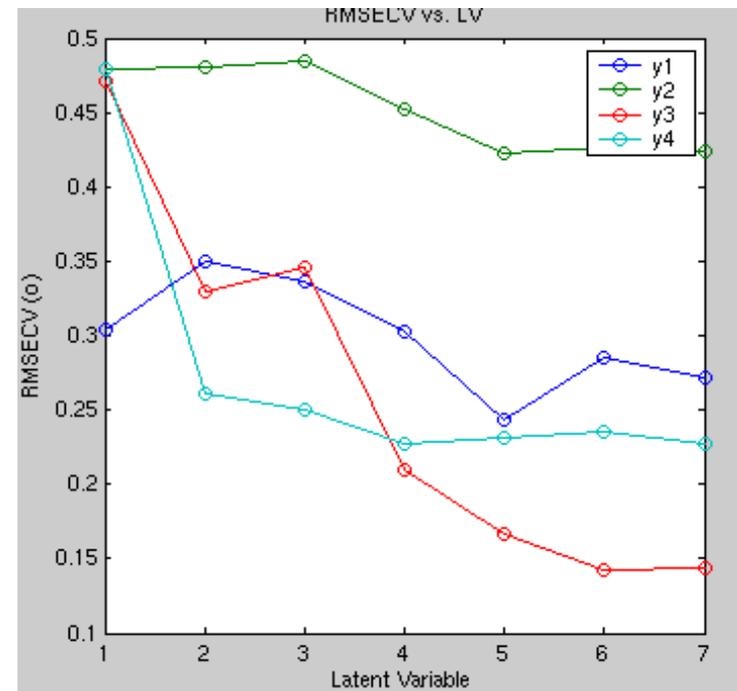
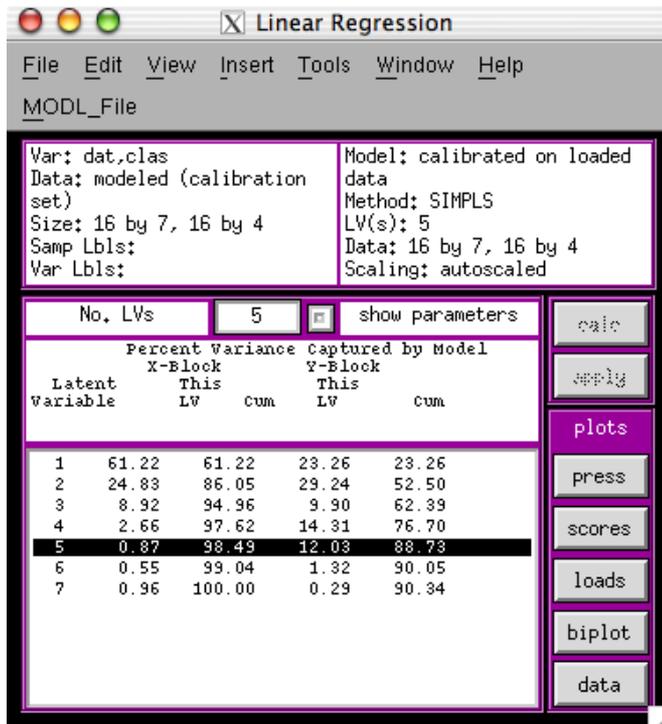
PLS-DA esempio

metodi fertilizzanti per mele

- Tre metodi di fertilizzante per la coltivazione delle mele
 - Urea, nitrati di calcio e potassio, ammonio e solfati
 - Più un controllo
- Quattro classi
- Ogni mela è caratterizzata da un pattern di sette features:
 - Azoto totale, azoto semi, fosforo, potassio, calcio, magnesio, peso

Y					X						
control	urea	potassium nitrates	ammonium and sulfate		total nitrogen	pit nitrogen	phosphorous	potassium	calcium	magnesium	weight
1	0	0	0	0	3240	1663	836	8747	218	388	97.3
1	0	0	0	0	3077	1663	891	8460	249	372	75.8
1	0	0	0	0	3205	1770	831	8575	261	376	78.5
1	0	0	0	0	3330	1755	889	8330	209	367	77.5
0	1	0	0	0	3755	1915	842	10375	145	408	108.2
0	1	0	0	0	5037	2180	930	10047	172	420	103.2
0	1	0	0	0	4753	2137	945	10447	160	421	95.3
0	1	0	0	0	4453	1967	850	9677	206	396	93.5
0	0	1	0	0	4200	2063	869	11190	184	398	111.8
0	0	1	0	0	5915	2050	1016	12060	184	461	109.7
0	0	1	0	0	5193	2210	958	11733	179	428	117.3
0	0	1	0	0	5347	2167	919	11910	191	420	99.6
0	0	0	1	0	5157	2357	1062	10210	136	401	86.7
0	0	0	1	1	7440	2975	1261	10820	122	446	85.5
0	0	0	1	1	6950	2527	1137	9710	191	408	70.8
0	0	0	1	1	4445	2075	846	8820	161	334	81.1

PLS-DA esempio metodi fertilizzanti per mele

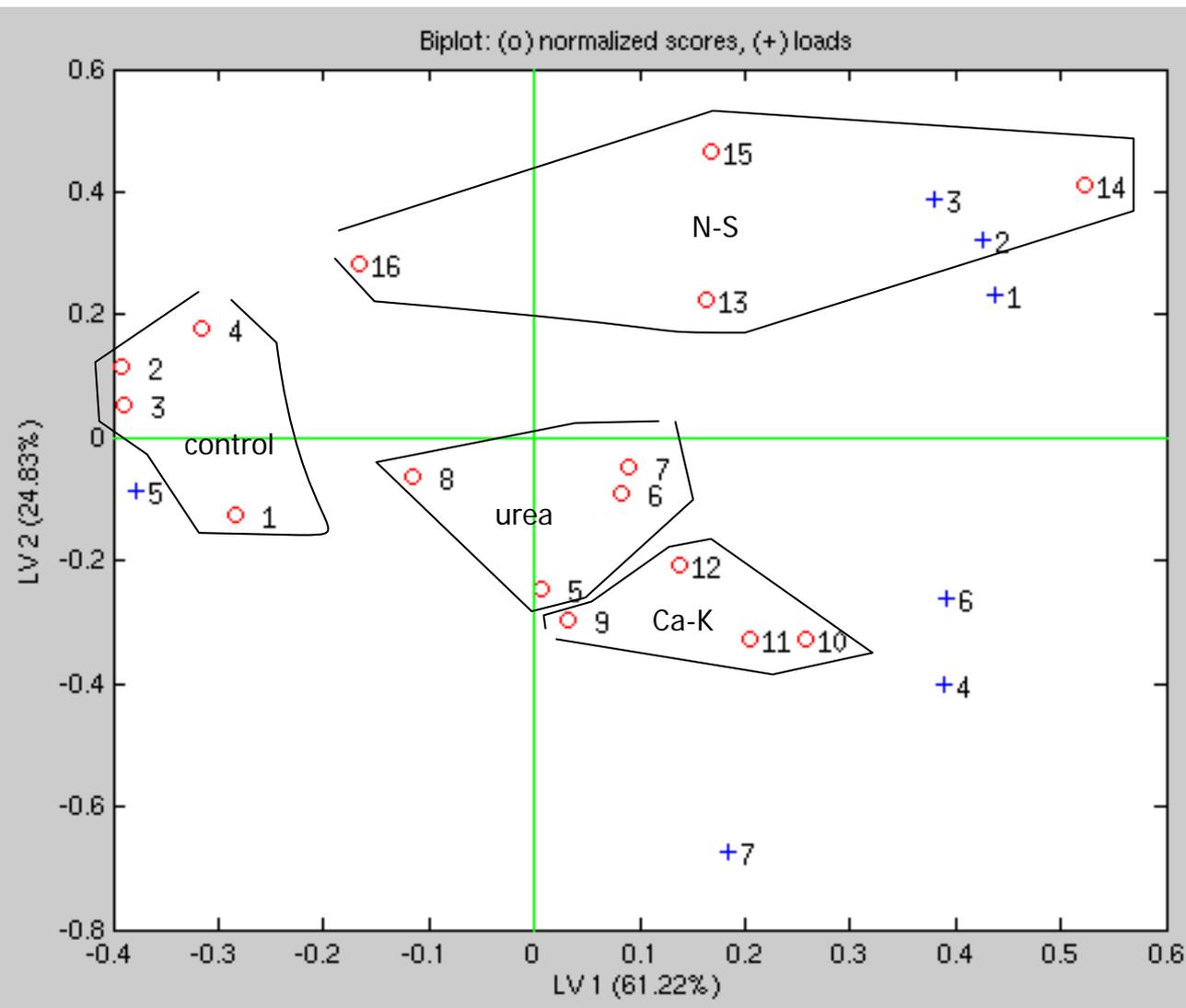


PLS-DA esempio

metodi fertilizzanti per mele

Y vero				Y stimato			
1	0	0	0	0.7839	0.4456	-0.0168	-0.2128
1	0	0	0	1.1728	-0.3482	0.1035	0.0718
1	0	0	0	0.8882	0.1623	0.0387	-0.0892
1	0	0	0	0.8729	0.0584	-0.2114	0.2801
0	1	0	0	0.0515	0.9332	0.0552	-0.0398
0	1	0	0	0.0116	0.9322	-0.0086	0.0648
0	1	0	0	0.0748	0.7485	0.0089	0.1678
0	1	0	0	0.1801	0.7226	0.0919	0.0053
0	0	1	0	0.0482	-0.1203	0.9887	0.0835
0	0	1	0	0.0820	0.2404	0.8671	-0.1895
0	0	1	0	-0.0390	-0.0669	1.0746	0.0313
0	0	1	0	-0.1771	0.0942	0.9599	0.1230
0	0	0	1	0.1673	-0.0846	0.1036	0.8137
0	0	0	1	-0.2136	0.1390	-0.0245	1.0990
0	0	0	1	0.1372	-0.0736	0.0300	0.9064
0	0	0	1	-0.0410	0.2174	-0.0608	0.8844

PLS-DA esempio metodi fertilizzanti per mele



Gli Scores mostrano:

- Separazione tra i quattro gruppi
- Due direzioni di allontanamento dal fertilizzante normale: NS e urea Ca-K

I Loadings mostrano:

- Il trattamento N-S aumenta l'azoto totale e nei semi ed il fosforo
- I trattamenti con urea e Ca-K aumentano potassio, magnesio ed il peso del frutto
- La quantità di calcio maggiore si trova nelle mele di controllo