

Méthodes à noyaux pour la détection de contexte.

Vers un fonctionnement autonome des méthodes à
noyaux

Gaëlle Loosli

gaelle.loosli@insa-rouen.fr

Laboratoire d'Informatique, de Traitement de l'Information et des
Systèmes.

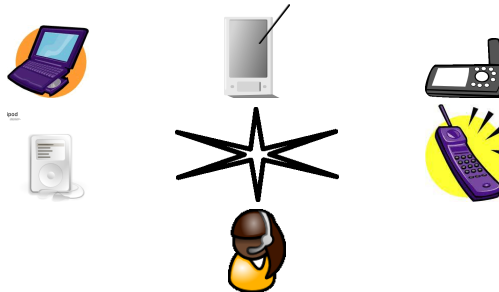
Directeur de thèse : Stéphane Canu

Soutenance de thèse

Saint Etienne du Rouvray - 27 octobre 2006



Détection de contexte



Tous ces appareils pourraient tirer parti du contexte de leur utilisation

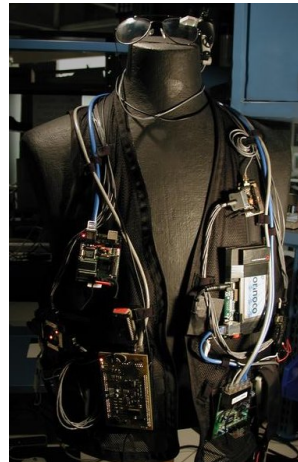
Détection de contexte

Capteurs

- accéléromètres
- pression sanguine
- conductivité
- respiration

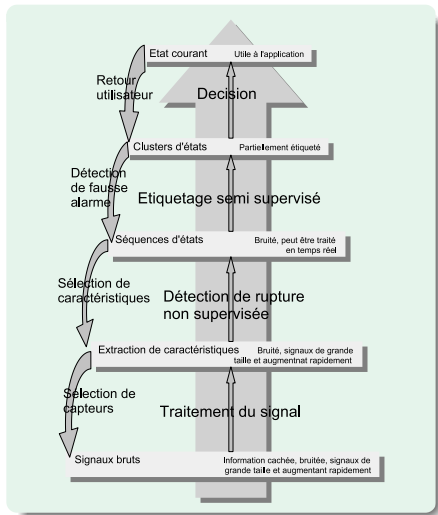
Signaux

- bruités
- non étiquetés
- variables selon l'utilisateur
- variables selon le moment



(media.mit.edu)

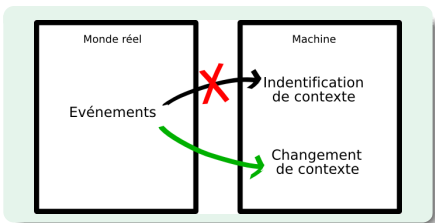
Architecture



Application cible

- 1 capteurs : ceux que l'utilisateur peut porter sur lui
- 2 classes : juste celles utiles à l'application [1]

Architecture

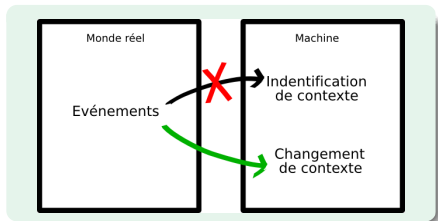


Application cible

- 1 capteurs : ceux que l'utilisateur peut porter sur lui
- 2 classes : juste celles utiles à l'application [1]

Ethique

Architecture



Application cible

- 1 capteurs : ceux que l'utilisateur peut porter sur lui
- 2 classes : juste celles utiles à l'application [1]

Ethique



Gaëlle Loosli, Sang-Goog Lee, Vincent Guigue, Alain Rakotomamonjy, and Stéphane Canu.
Perception d'états affectifs et apprentissage.
Revue d'intelligence artificielle, Edition spéciale Interaction Emotionnelle, 20 No.4-5 :553-582,
2006.

Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome
- 4 Conclusion

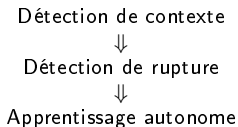
Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome
- 4 Conclusion

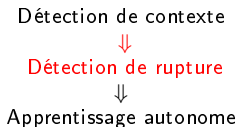
Détection de contexte
↓
Détection de rupture

Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome
- 4 Conclusion



- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome
- 4 Conclusion



Détection de rupture

Signaux

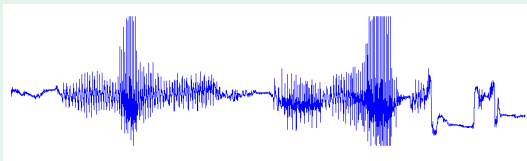
Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé

Détection de rupture

Signaux



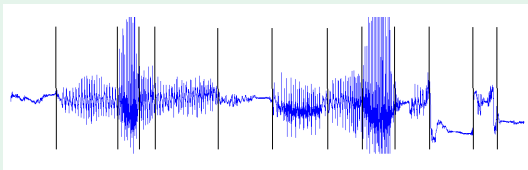
Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé

Détection de rupture

Signaux



- 1 il y a-t-il eu rupture ?
- 2 quand il y a-t-il eu rupture ?
- 3 minimiser le retard de détection

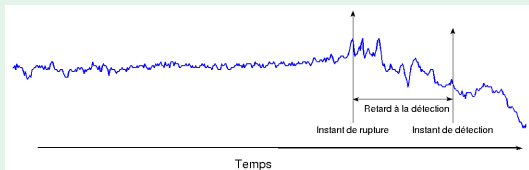
Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé

Détection de rupture

Signaux



- 1 il y a-t-il eu rupture ?
- 2 quand il y a-t-il eu rupture ?
- 3 minimiser le retard de détection

Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé

Détection de rupture

Signaux

Approches classiques [Basseville & Nikiforov, 1993]

- 1 utilisent un modèle
- 2 détecteurs dédiés (moyenne, variance,...)

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé

Détection de rupture

Signaux

Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

- 1 méthode à noyau (robuste)
- 2 méthode sans modèle

Principe utilisé

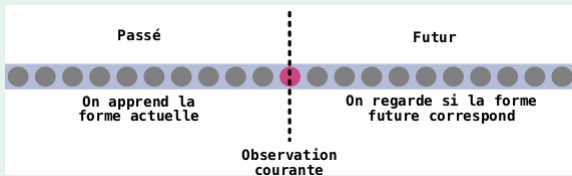
Détection de rupture

Signaux

Approches classiques [Basseville & Nikiforov, 1993]

Approche choisie : SVM à une classe [Desobry et al., 2005]

Principe utilisé



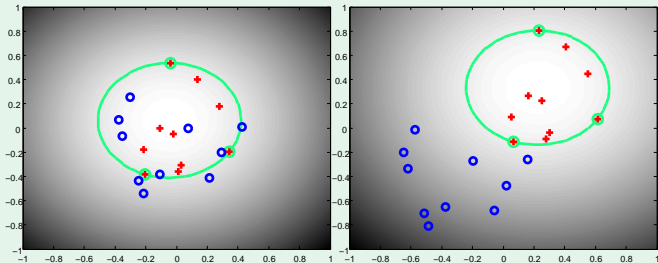
SVM à une classe

Apprendre avec des exemples d'une seule classe

- objectif : tracer le contour de la classe

- fonction de décision : $f(x_j) = \sum_{i=1}^m \alpha_i k(x_i, x_j)$

- le noyau : $k(x_i, \cdot)$ zone d'influence de x_i . $k(x_i, x_j) = \exp \frac{-(x_i - x_j)^2}{\sigma^2}$



Test d'hypothèse

Hypothèses

- hypothèse 0 : pas de changement de contexte à l'instant considéré (par défaut)
- hypothèse 1 : changement de contexte à l'instant considéré
- passé futur

Test [Canu & Smola, 2006]

- on montre que sous certaines conditions, la région d'acceptation du test est définie sous la forme des sorties du SVM à une classe.
- région d'acceptation de l'hypothèse 0 : $-\sum_{j=t+1}^n \sum_{i=1}^t \alpha_i k(x_j, x_i) < \text{seuil}$
- il reste à déterminer le seuil

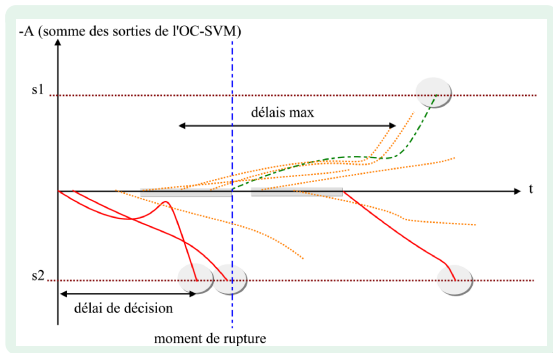
Retard à la détection

Décision

Quand prendre une décision ?

Seuils

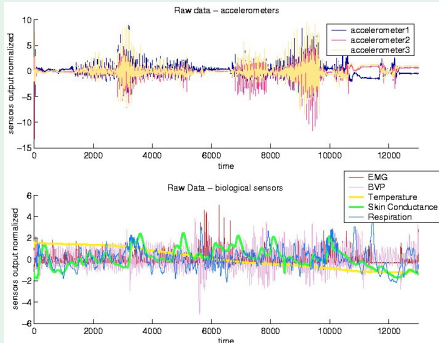
- $\Sigma \geq s1$:
changement
- $\Sigma \leq s2$: pas de
changement
- $s2 < \Sigma < s1$:
regarder un
nouveau point



Application à des données réelles

Acquisition de données

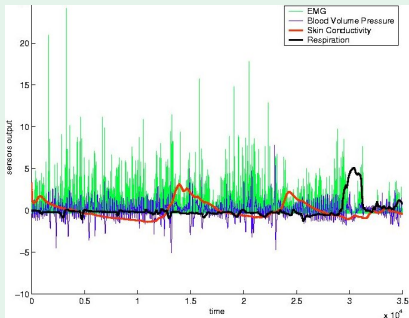
- utilisateur équipé de capteurs physiologiques et d'accéléromètres
- se déplace librement en étant filmé



Application à des données réelles

Acquisition de données

- utilisateur équipé de capteurs physiologiques
- joue à Xblast



Résultats

Déplacements

A posteriori	Obs.	Auto
Nb de ruptures	16	17
Détection	16	15 (94%)
Fausse alarme	0	2 (12%)
Non Détection	1	2 (13%)

Jeu vidéo

Joueur	Obs.	Auto
50	39	55
50	39	36 - 32 (72% - 82%)
0	0	19 - 23 (35% - 42%)
0	11	14 - 7 (22% - 18%)

Remarques

- l'étiquetage pose un problème
- les accéléromètres sont d'une grande utilité pour la détection de changement
- filtrage d'alarme par l'intégration dans le système global

Conclusion

Cette étude de faisabilité montre que la détection de rupture peut permettre la détection de contexte.

Les SVM sont efficaces mais ne peuvent pas directement être utilisés en ligne et pour de grandes masses de données.



Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.
Context changes detection by one-class svms.
UM 2005 : Workshop on Machine Learning for User Modeling : Challenges, 2005.



Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.
Context retrieval by rupture detection.
Conférence d'Apprentissage, 2005.



Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.
Rupture detection for context aware applications.
CEUR Workshop, 2005.

Conclusion

Cette étude de faisabilité montre que la détection de rupture peut permettre la détection de contexte.

Les SVM sont efficaces mais ne peuvent pas directement être utilisés en ligne et pour de grandes masses de données.



Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.

Context changes detection by one-class svms.

UM 2005 : Workshop on Machine Learning for User Modeling : Challenges, 2005.



Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.

Context retrieval by rupture detection.

Conférence d'Apprentissage, 2005.



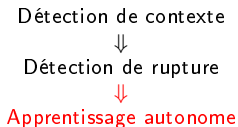
Gaëlle Loosli, Sang-Goog Lee, and Stéphane Canu.

Rupture detection for context aware applications.

CEUR Workshop, 2005.

Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome
- 4 Conclusion



Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome**
 - Méthode d'apprentissage : SVM
 - Apprentissage en ligne
 - Réglage des hyper-paramètres
- 4 Conclusion

Méthode d'apprentissage : SVM

La classification

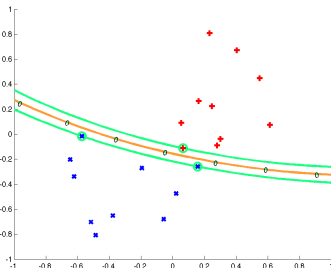
On considère un problème de classification binaire, avec les exemples d'apprentissage $x_1 \dots x_m \in \mathcal{X}$ et les classes $y_1 \dots y_m \in \{+1, -1\}$.

Contraintes strictes

$$\begin{cases} \min_{f,b} \frac{1}{2} \|f\|^2 \\ y_i (f(x_i) + b) \geq 1 \quad \forall i \in [1, \dots, m] \end{cases}$$

$$\begin{cases} \max_{\alpha} -\frac{1}{2} \alpha^\top G \alpha \\ \alpha^\top \mathbf{y} = 0 \\ 0 \leq \alpha_i \end{cases} \quad \forall i \in [1, \dots, m]$$

avec $G(i, j) = \frac{1}{m} y_i y_j k(x_i, x_j)$.



Méthode d'apprentissage : SVM

La classification

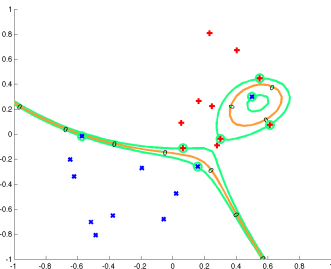
On considère un problème de classification binaire, avec les exemples d'apprentissage $x_1 \dots x_m \in \mathcal{X}$ et les classes $y_1 \dots y_m \in \{+1, -1\}$.

Contraintes strictes

$$\begin{cases} \min_{f,b} \frac{1}{2} \|f\|^2 \\ y_i (f(x_i) + b) \geq 1 \quad \forall i \in [1, \dots, m] \end{cases}$$

$$\begin{cases} \max_{\alpha} -\frac{1}{2} \alpha^\top G \alpha \\ \alpha^\top y = 0 \\ 0 \leq \alpha_i \end{cases} \quad \forall i \in [1, \dots, m]$$

avec $G(i, j) = \frac{1}{m} y_i y_j k(x_i, x_j)$.



Méthode d'apprentissage : SVM

La classification

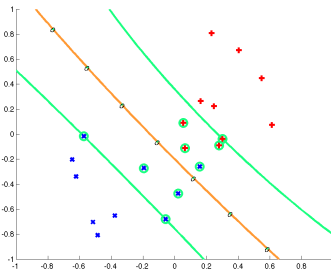
On considère un problème de classification binaire, avec les exemples d'apprentissage $x_1 \dots x_m \in \mathcal{X}$ et les classes $y_1 \dots y_m \in \{+1, -1\}$.

Contraintes douces

$$\begin{cases} \min_{f, b, \xi} \frac{1}{2} \|f\|^2 + C \sum_{i=1}^m \xi_i \\ y_i (f(x_i) + b) \geq 1 - \xi_i & \forall i \in [1, \dots, m] \\ \xi_i \geq 0 & \forall i \in [1, \dots, m] \end{cases}$$

$$\begin{cases} \max_{\alpha} -\frac{1}{2} \alpha^\top G \alpha - \alpha^\top \mathbf{1} \\ \alpha^\top \mathbf{y} = 0 \\ 0 \leq \alpha_i \leq C & \forall i \in [1, \dots, m] \end{cases}$$

avec $G(i, j) = \frac{1}{m} y_i y_j k(x_i, x_j)$.



Répartition des points

Description

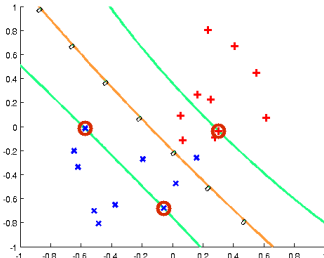
- situés sur la marge, contribution *libre* à la solution
- situés du mauvais côté de la marge, contribution *limitée* à la solution
- situés du bon côté de la marge, contribution *nulle* à la solution

Caractéristiques

- $y_i(f(x_i) + b) = 1$ et $0 < \alpha_i < C$
- $y_i(f(x_i) + b) < 1$ et $\alpha_i = C$
- $y_i(f(x_i) + b) > 1$ et $\alpha_i = 0$

Groupes de points

- Groupe I_a
- Groupe I_C
- Groupe I_0

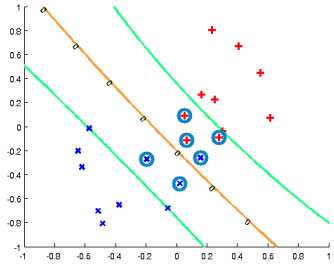


Répartition des points

- ### Description
- situés sur la marge, contribution *libre* à la solution
 - situés du mauvais côté de la marge, contribution *limitée* à la solution
 - situés du bon côté de la marge, contribution *nulle* à la solution

- ### Groupes de points
- Groupe I_a
 - Groupe I_C
 - Groupe I_0

- ### Caractéristiques
- $y_i(f(x_i) + b) = 1$ et $0 < \alpha_i < C$
 - $y_i(f(x_i) + b) < 1$ et $\alpha_i = C$
 - $y_i(f(x_i) + b) > 1$ et $\alpha_i = 0$

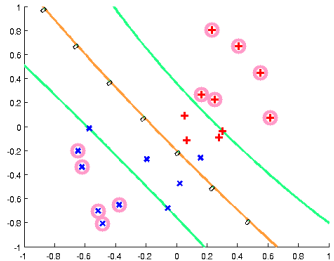


Répartition des points

- ### Description
- situés sur la marge, contribution *libre* à la solution
 - situés du mauvais côté de la marge, contribution *limitée* à la solution
 - situés du bon côté de la marge, contribution *nulle* à la solution

- ### Groupes de points
- Groupe I_a
 - Groupe I_C
 - Groupe I_0

- ### Caractéristiques
- $y_i(f(x_i) + b) = 1$ et $0 < \alpha_i < C$
 - $y_i(f(x_i) + b) < 1$ et $\alpha_i = C$
 - $y_i(f(x_i) + b) > 1$ et $\alpha_i = 0$

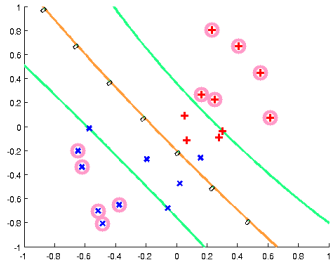


Répartition des points

- ### Description
- situés sur la marge, contribution *libre* à la solution
 - situés du mauvais côté de la marge, contribution *limitée* à la solution
 - situés du bon côté de la marge, contribution *nulle* à la solution

- ### Caractéristiques
- $y_i(f(x_i) + b) = 1$ et $0 < \alpha_i < C$
 - $y_i(f(x_i) + b) < 1$ et $\alpha_i = C$
 - $y_i(f(x_i) + b) > 1$ et $\alpha_i = 0$

- ### Groupes de points
- Groupe I_a
 - Groupe I_C pas de calcul de α
 - Groupe I_0 pas de calcul de α



SimpleSVM - Contraintes actives

Algorithme

```

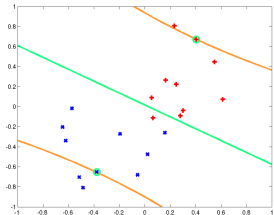
Initialiser
tant que points mal classés faire
  calculer  $\alpha_{I_a}$ 
  si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
    retirer un point de  $I_a$ 
  sinon
    si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
      alors
        ajouter un point à  $I_a$ 
    fin si
  fin si
fin tant que
  
```

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
  calculer  $\alpha_{I_a}$ 
  si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
    retirer un point de  $I_a$ 
  sinon
    si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
      alors
         $I_C$  ou  $I_0$  vers  $I_a$ 
      fin si
    fin si
  fin tant que
  
```



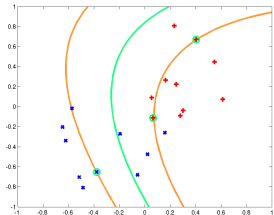
$$\alpha_{I_a} = \{4.7504; 4.7504\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
  calculer  $\alpha_{I_a}$ 
  si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
     $I_a$  vers  $I_C$  ou  $I_0$ 
  sinon
    si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
      alors
        ajouter un point à  $I_a$ 
      fin si
    fin si
  fin tant que
  
```



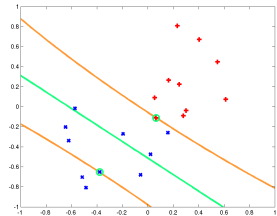
$$\alpha_{I_a} = \{58.2821; -38.3212; 96.6033\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
    calculer  $\alpha_{I_a}$ 
    si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
        retirer un point de  $I_a$ 
    sinon
        si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
            alors
                 $I_C$  ou  $I_0$  vers  $I_a$ 
            fin si
        fin si
    fin tant que
    
```



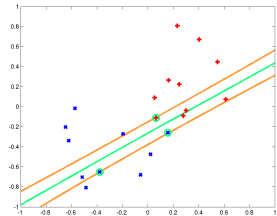
$$\alpha_{I_a} = \{21.0157; 21.0157\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
    calculer  $\alpha_{I_a}$ 
    si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
         $I_a$  vers  $I_C$  ou  $I_0$ 
    sinon
        si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
        alors
            ajouter un point à  $I_a$ 
        fin si
    fin si
fin tant que
    
```



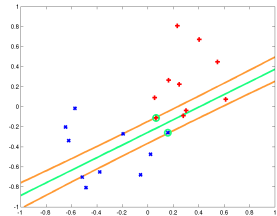
$$\alpha_{I_a} = \{-5.5362; 343.3578; 337.8216\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
    calculer  $\alpha_{I_a}$ 
    si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
        retirer un point de  $I_a$ 
    sinon
        si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
            alors
                 $I_C$  ou  $I_0$  vers  $I_a$ 
            fin si
        fin si
    fin tant que
    
```



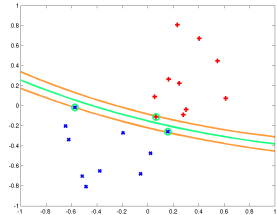
$$\alpha_{I_a} = \{336.5028; 336.5028\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
    calculer  $\alpha_{I_a}$ 
    si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
        retirer un point de  $I_a$ 
    sinon
        si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
            ajouter un point à  $I_a$ 
        fin si
    fin si
fin tant que
    
```



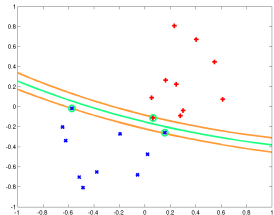
$$\alpha_{I_a} = \{641.3429; 773.3408; 131.9979\}$$

SimpleSVM - Contraintes actives

Algorithme

```

Initialiser
tant que points mal classés faire
    calculer  $\alpha_{I_a}$ 
    si  $\alpha_{I_a} \leq 0$  ou  $\alpha_{I_a} \geq C$  alors
        retirer un point de  $I_a$ 
    sinon
        si  $y(f(x_{I_0}) + b) \leq 1$  ou  $y(f(x_{I_C}) + b) \geq 1$ 
            alors
                ajouter un point à  $I_a$ 
            fin si
        fin si
    fin tant que
    
```



Gaëlle Loosli.

Fast svm toolbox in Matlab based on SimpleSVM algorithm, 2004.

<http://asi.insa-rouen.fr/~gloosli/simpleSVM.html>.



Gaëlle Loosli, Stéphane Canu, S.V.N. Vishwanathan, Alexander J. Smola, and Monojit Chattopadhyay.

Boîte à outils svm simple et rapide.

Revue d'intelligence artificielle, 19, 2005.

Méthodes de résolution

Résolution exacte

- Point Intérieur
- SimpleSVM
- SMO

Résolution approchée

- CVM
- SVM

Méthodes de résolution

Résolution exacte

- Point Intérieur
- SimpleSVM
- SMO

Résolution approchée

- CVM
- LASVM

Méthodes de résolution

Résolution exacte

- Point Intérieur
- SimpleSVM
- SMO

Résolution approchée

- CVM
- LASVM

Méthodes de résolution

Résolution exacte

- Point Intérieur
- SimpleSVM
- SMO

Résolution approchée

- CVM
- LASVM

Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome**
 - Méthode d'apprentissage : SVM
 - Apprentissage en ligne
 - Réglage des hyper-paramètres
- 4 Conclusion

Apprentissage en ligne

En ligne

- de nouveaux points sont donnés au fur et à mesure de l'apprentissage
- arrêt de l'apprentissage quand il n'y a plus de nouveaux points

Comparaison

	SimpleSVM	LASVM [Bordes et al., 2005]
A chaque point		
Finalisation		

Apprentissage en ligne

En ligne

- de nouveaux points sont donnés au fur et à mesure de l'apprentissage
- arrêt de l'apprentissage quand il n'y a plus de nouveaux points

Comparaison

	SimpleSVM	LASVM [Bordes et al., 2005]
A chaque point	calcul des α sur I_a puis vérification d'optimalité pour tous les points connus	calcul des α des couples de points vecteurs supports (le <i>Process</i> et <i>Re-process</i>)
Finalisation		

Apprentissage en ligne

En ligne

- de nouveaux points sont donnés au fur et à mesure de l'apprentissage
- arrêt de l'apprentissage quand il n'y a plus de nouveaux points

Comparaison

	SimpleSVM	LASVM [Bordes et al., 2005]
A chaque point	calcul des α sur I_a puis vérification d'optimalité pour tous les points connus	calcul des α des couples de points vecteurs supports (le <i>Process</i> et <i>Re-process</i>)
Finalisation	-	optimisation complète sur $\{I_a, I_C\}$

Application à grande échelle

Quelles sont les capacités de LASVM ?

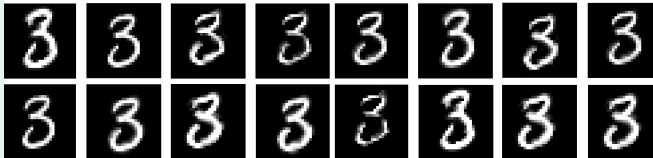
- choix d'un problème *extensible*
- chaque nouvel exemple est généré au fur et à mesure
- le problème des invariances est un bon candidat pour cette expérience

Application à grande échelle

Quelles sont les capacités de LASVM ?

- choix d'un problème *extensible*
- chaque nouvel exemple est généré au fur et à mesure
- le problème des invariances est un bon candidat pour cette expérience

Il est possible de calculer rapidement des déformations d'images



Application à grande échelle

Démarche expérimentale

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

Résultats

Conclusion

Application à grande échelle

Démarche expérimentale

- apprentissage lancé sur la base d'origine
- puis l'apprentissage continue sur des déformations des exemples originaux
- continuer tant que la mémoire et le temps disponibles le permettent

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

Résultats

Conclusion

Application à grande échelle

Démarche expérimentale

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

- 10 classes (de 0 à 9) : apprentissage de dix classifieurs type 1 contre tous
- nombre d'exemple générés : 8 millions
- soit 80 millions de points traités
- nombre de processeurs : 1
- mémoire : 8 Go

Résultats

Conclusion

Application à grande échelle

Démarche expérimentale

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

Résultats

- temps d'apprentissage et de test : 8 jours
- nombre de vecteurs supports : 120 000 (soit 0.15% des exemples vus)
- performance : 0.7% d'erreur (réseaux à convolution : 0.4% méthodes à noyaux : 0.7%/0.8%)

Conclusion

Application à grande échelle

Démarche expérimentale

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

Résultats

Conclusion

- traiter une si grande base de données n'est pas envisageable par une méthode à noyau hors ligne (problèmes de mémoire)
- c'est le plus grand SVM reporté à ce jour sur un seul processeur

Application à grande échelle

Démarche expérimentale

Base de données de chiffres manuscrits MNIST [LeCun et al., 1998]

Résultats

Conclusion



Gaëlle Loosli, Stéphane Canu, and Léon Bottou.
Svm et apprentissage des très grandes bases de données.
CAp - Conférence d'apprentissage, 2006.



Gaëlle Loosli, Stéphane Canu, SVN Vishwanathan, and Alexander J.Smola.
Invariances in classification : an efficient svm implementation.
Applied Stochastic Models and Data Analysis, 2005.

Plan

- 1 Introduction
- 2 Détection de rupture
- 3 Apprentissage autonome**
 - Méthode d'apprentissage : SVM
 - Apprentissage en ligne
 - Réglage des hyper-paramètres
- 4 Conclusion

Interprétation de l'hyper-paramètre C

Contraintes dans le dual : $0 \leq \alpha_j \leq C$

Réglage du relâchement des contraintes : C

- contraintes strictes : $C \rightarrow \infty$
- contraintes fortes : C *grand*
- contraintes très relâchées : C *petit*

ν -SVM

ν -SVM est une reformulation des SVM où ν joue le rôle de C avec :

- $1 \leq \nu \leq n$ (bornes connues)
- ν indique le nombre minimum de points dans la solution
- $0 \leq \alpha \leq 1$
- $\nu = 1 \Leftrightarrow C \rightarrow \infty$ et $\nu = n \Leftrightarrow C \rightarrow 0$

Interprétation de l'hyper-paramètre C

Contraintes dans le dual : $0 \leq \alpha_j \leq C$

Réglage du relâchement des contraintes : C

- contraintes strictes : $C \rightarrow \infty$
- contraintes fortes : C *grand*
- contraintes très relâchées : C *petit*

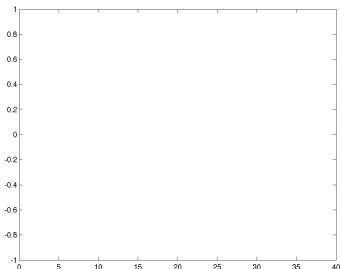
ν -SVM

ν -SVM est une reformulation des SVM où ν joue le rôle de C avec :

- $1 \leq \nu \leq n$ (bornes connues)
- ν indique le nombre minimum de points dans la solution
- $0 \leq \alpha \leq 1$
- $\nu = 1 \Leftrightarrow C \rightarrow \infty$ et $\nu = n \Leftrightarrow C \rightarrow 0$

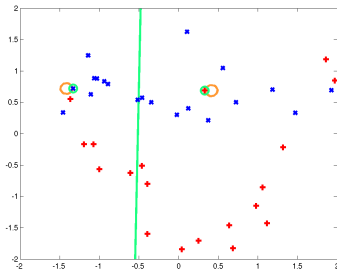
Chemin de régularisation

Suivi des $y_i \alpha_j$



$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

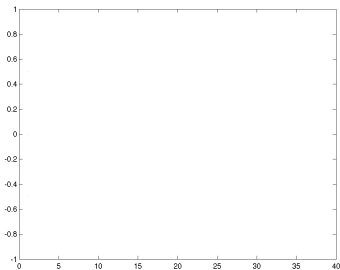
Suivi du chemin de régularisation



Recherche de la solution sans erreur
d'apprentissage

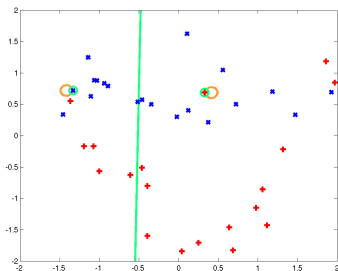
Chemin de régularisation

Suivi des $y_j \alpha_j$



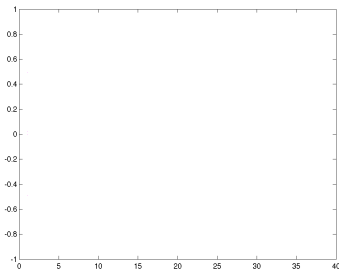
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



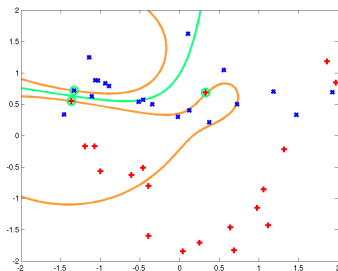
Chemin de régularisation

Suivi des $y_j \alpha_j$



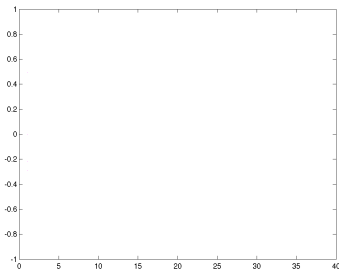
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



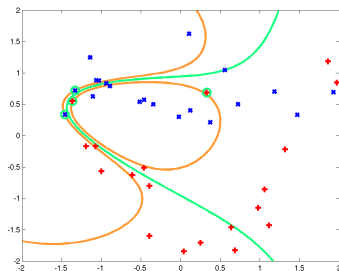
Chemin de régularisation

Suivi des $y_j \alpha_j$



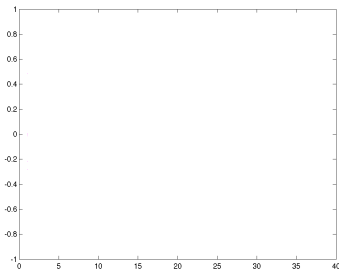
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



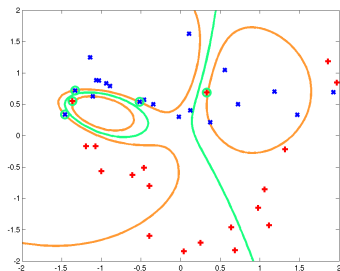
Chemin de régularisation

Suivi des $y_i \alpha_j$



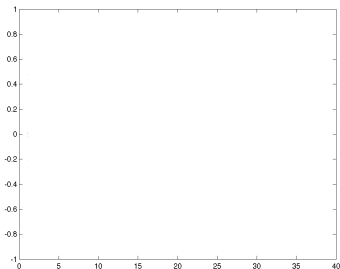
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



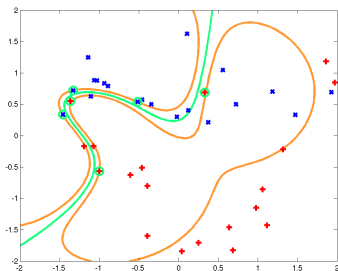
Chemin de régularisation

Suivi des $y_i \alpha_j$



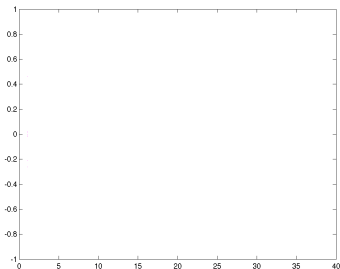
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



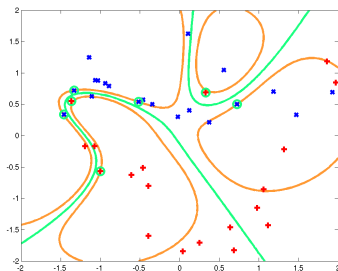
Chemin de régularisation

Suivi des $y_i \alpha_j$



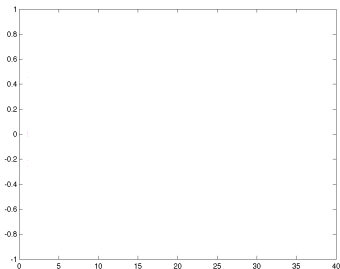
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



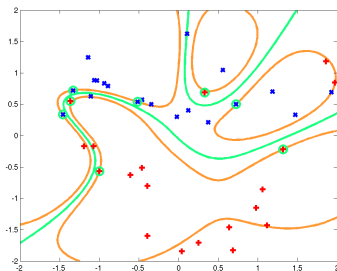
Chemin de régularisation

Suivi des $y_i \alpha_i$



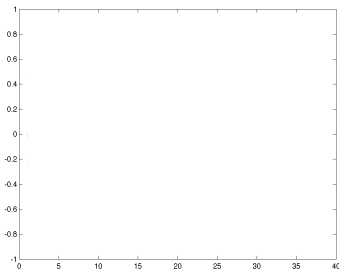
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



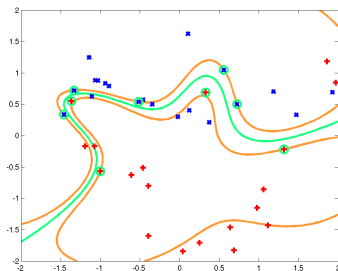
Chemin de régularisation

Suivi des $y_i \alpha_j$



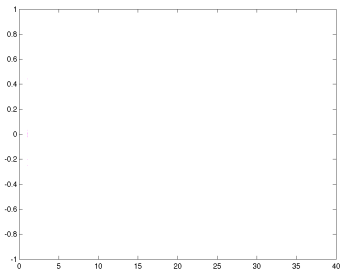
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



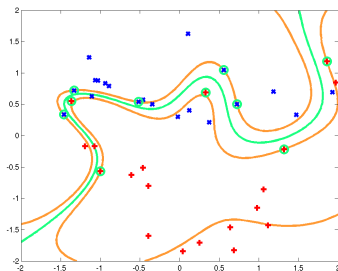
Chemin de régularisation

Suivi des $y_i \alpha_i$



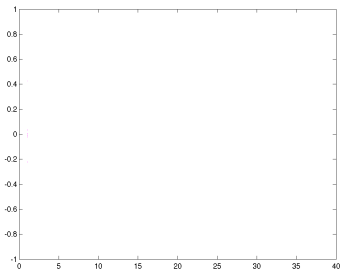
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



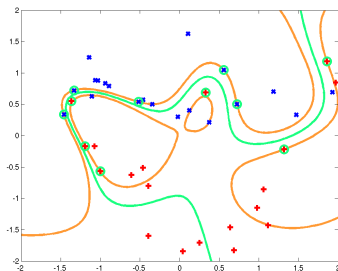
Chemin de régularisation

Suivi des $y_j \alpha_j$



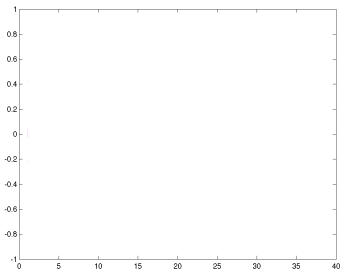
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



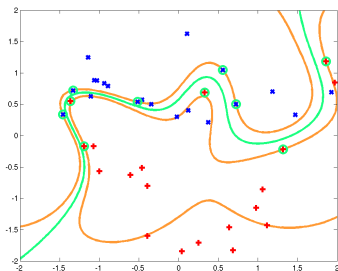
Chemin de régularisation

Suivi des $y_j \alpha_j$



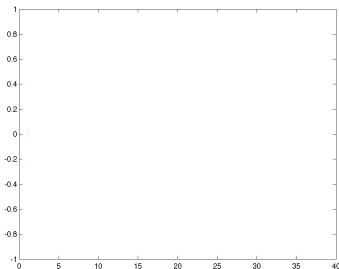
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



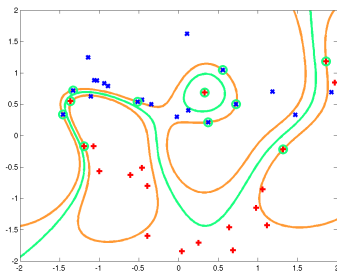
Chemin de régularisation

Suivi des $y_i \alpha_i$



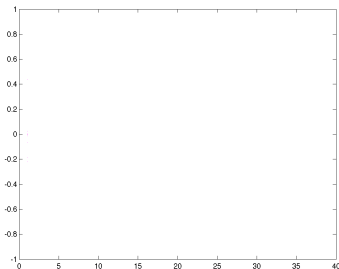
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



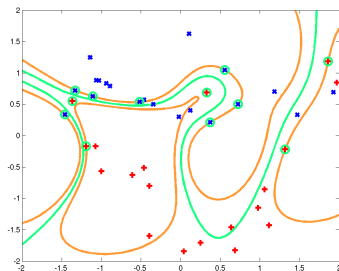
Chemin de régularisation

Suivi des $y_j \alpha_j$



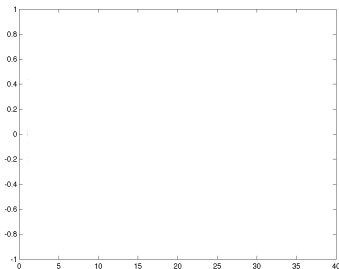
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



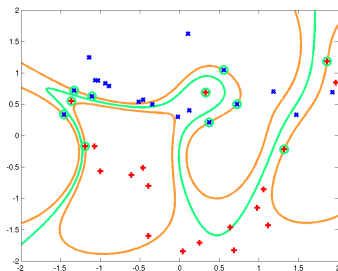
Chemin de régularisation

Suivi des $y_i \alpha_i$



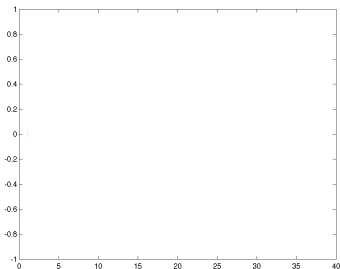
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



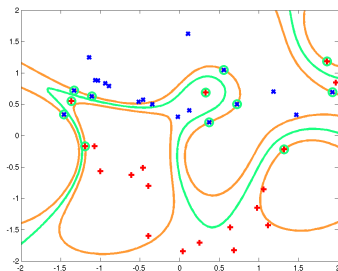
Chemin de régularisation

Suivi des $y_j \alpha_j$



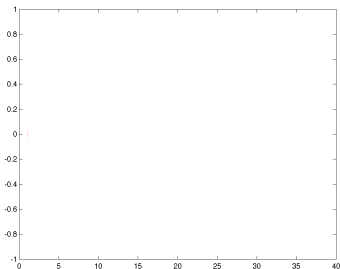
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



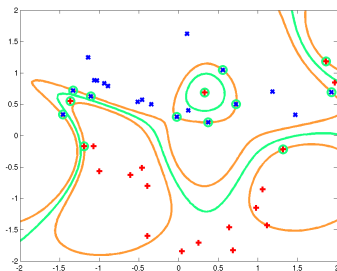
Chemin de régularisation

Suivi des $y_j \alpha_j$



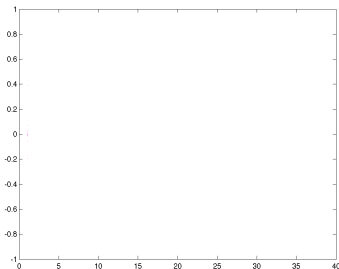
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



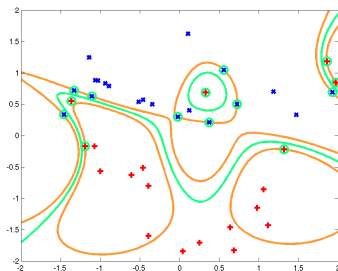
Chemin de régularisation

Suivi des $y_j \alpha_j$



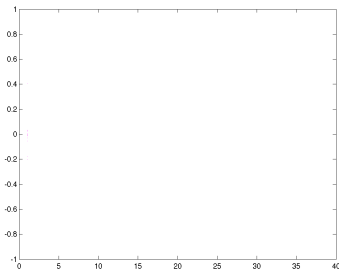
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



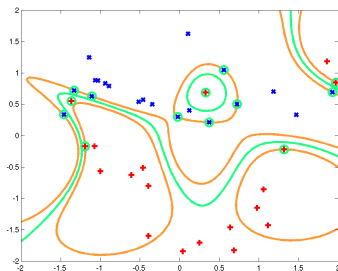
Chemin de régularisation

Suivi des $y_i \alpha_i$



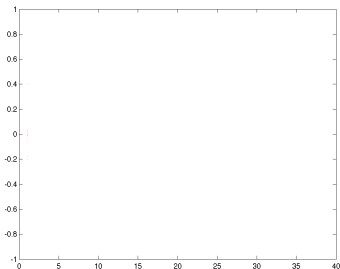
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



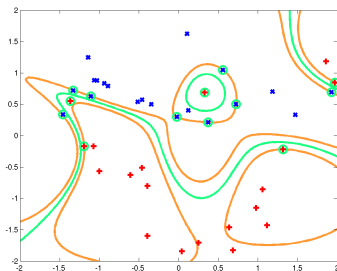
Chemin de régularisation

Suivi des $y_j \alpha_j$



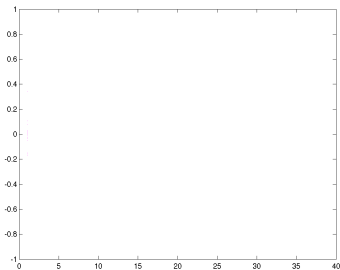
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



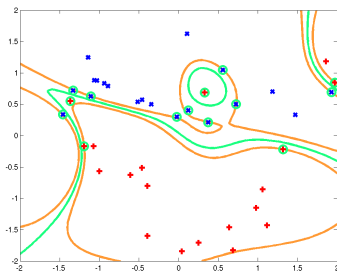
Chemin de régularisation

Suivi des $y_i \alpha_i$



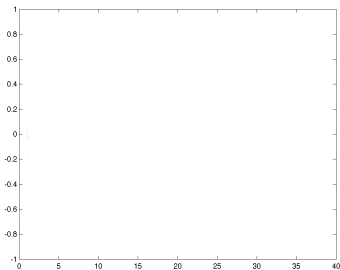
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



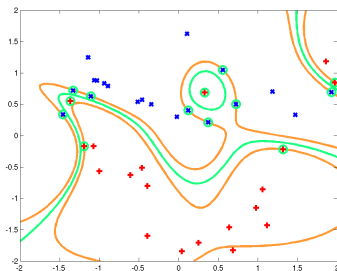
Chemin de régularisation

Suivi des $y_j \alpha_j$



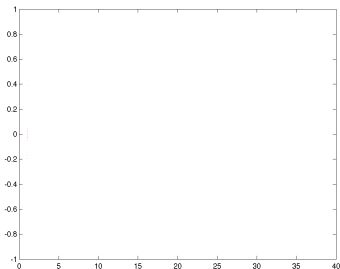
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



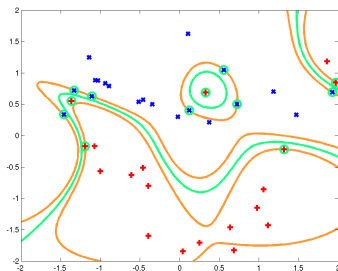
Chemin de régularisation

Suivi des $y_i \alpha_i$



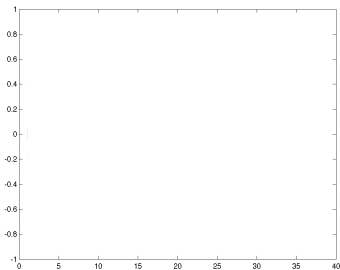
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



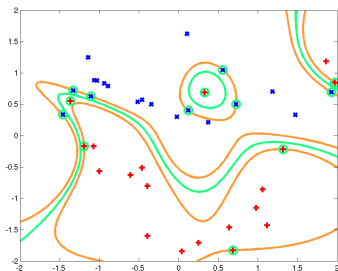
Chemin de régularisation

Suivi des $y_i \alpha_i$



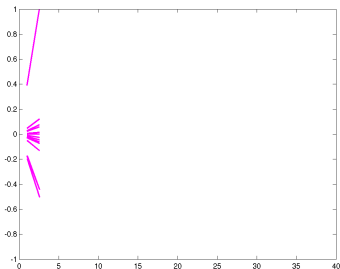
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



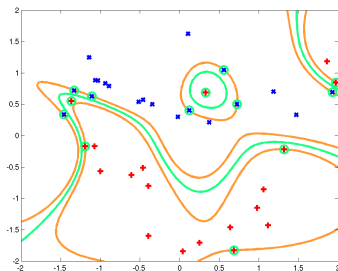
Chemin de régularisation

Suivi des $y_i \alpha_i$



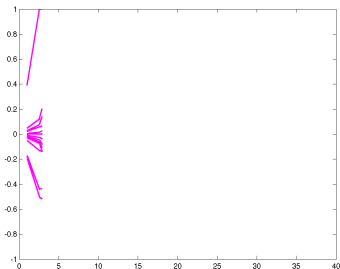
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



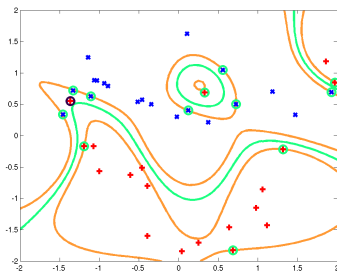
Chemin de régularisation

Suivi des $y_j \alpha_j$



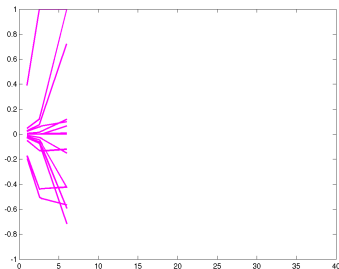
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



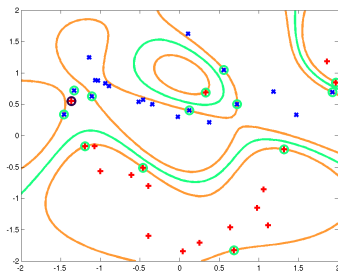
Chemin de régularisation

Suivi des $y_i \alpha_i$



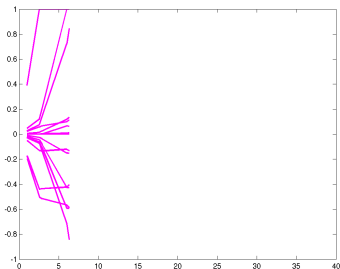
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



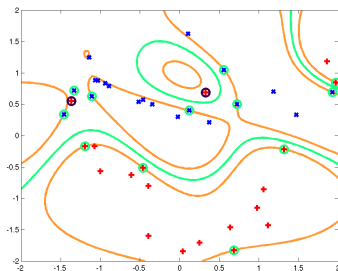
Chemin de régularisation

Suivi des $y_j \alpha_j$



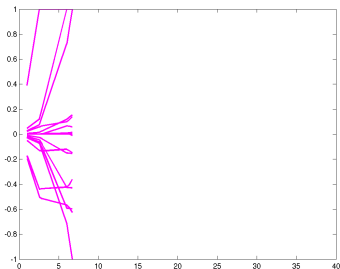
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



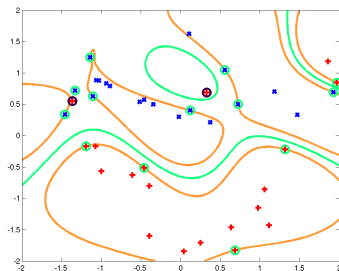
Chemin de régularisation

Suivi des $y_i \alpha_i$



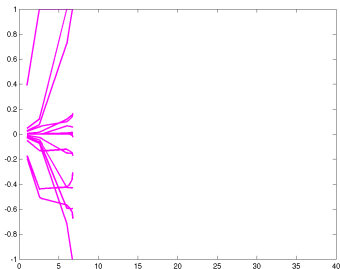
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



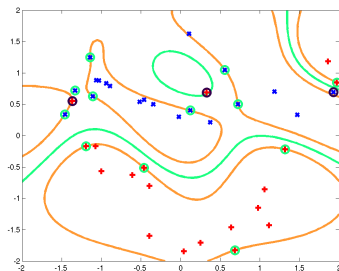
Chemin de régularisation

Suivi des $y_i \alpha_i$



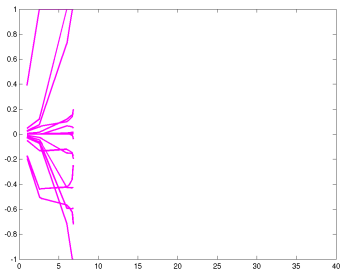
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



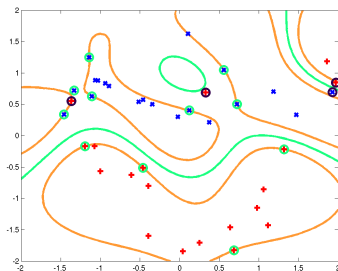
Chemin de régularisation

Suivi des $y_j \alpha_j$



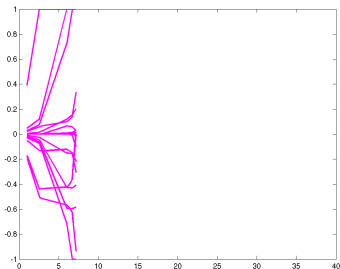
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



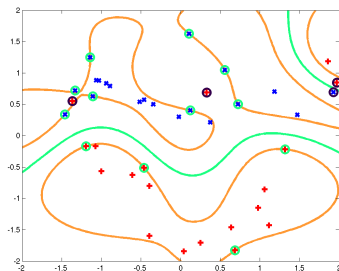
Chemin de régularisation

Suivi des $y_i \alpha_i$



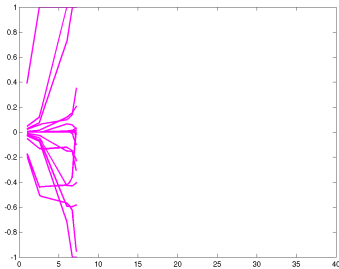
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



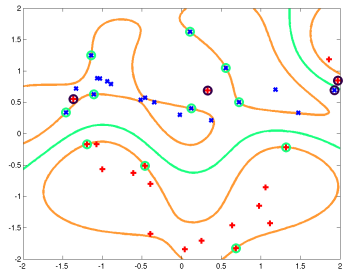
Chemin de régularisation

Suivi des $y_j \alpha_j$



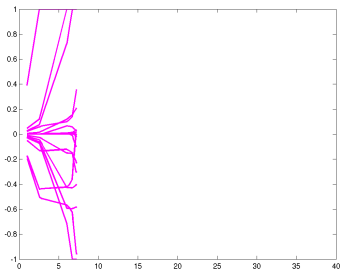
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



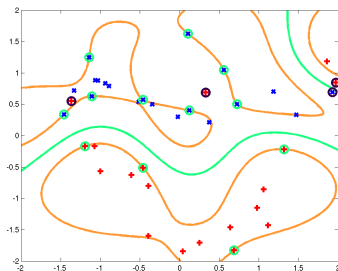
Chemin de régularisation

Suivi des $y_i \alpha_i$



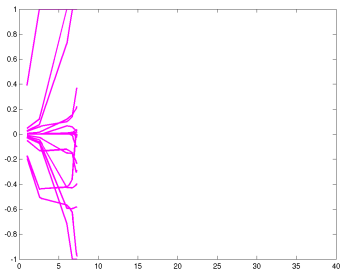
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



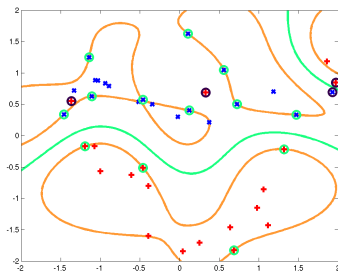
Chemin de régularisation

Suivi des $y_j \alpha_j$



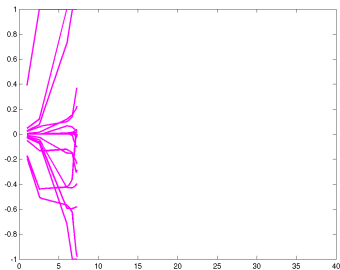
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



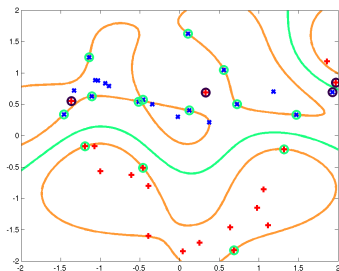
Chemin de régularisation

Suivi des $y_j \alpha_j$



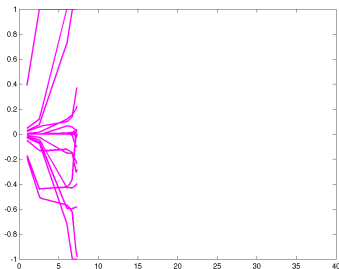
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



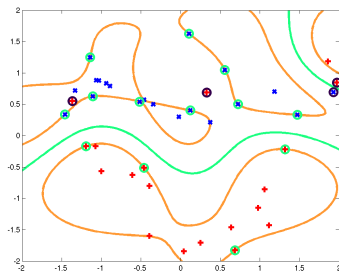
Chemin de régularisation

Suivi des $y_j \alpha_j$



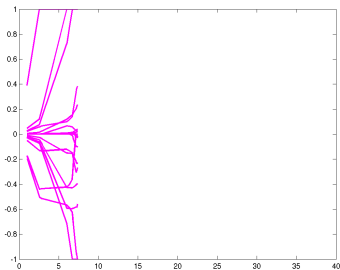
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



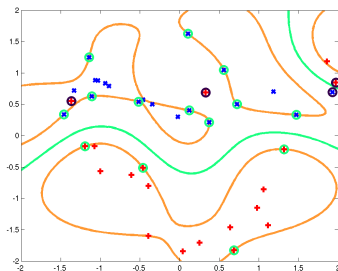
Chemin de régularisation

Suivi des $y_j \alpha_j$



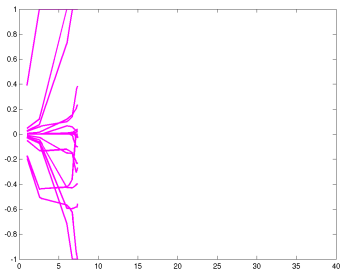
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



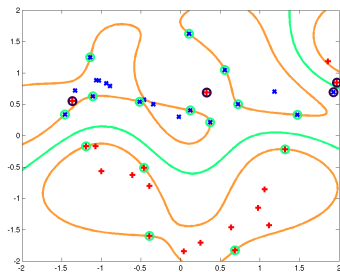
Chemin de régularisation

Suivi des $y_j \alpha_j$



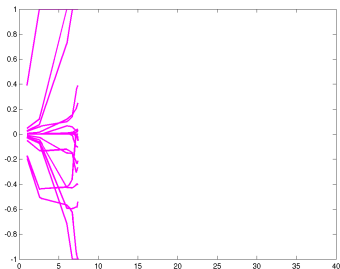
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



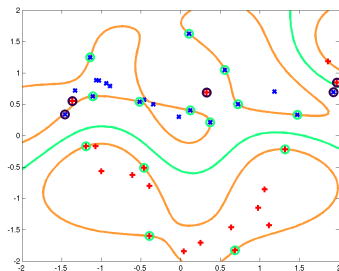
Chemin de régularisation

Suivi des $y_i \alpha_i$



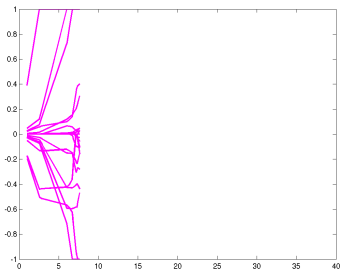
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



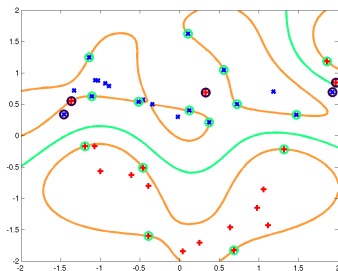
Chemin de régularisation

Suivi des $y_j \alpha_j$



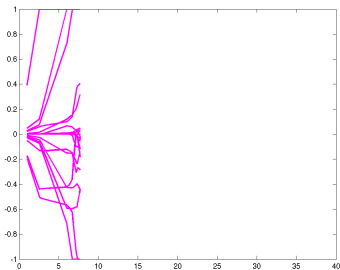
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



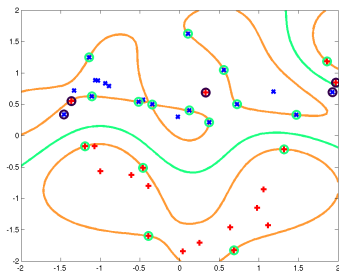
Chemin de régularisation

Suivi des $y_j \alpha_j$



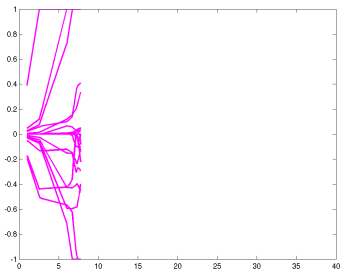
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



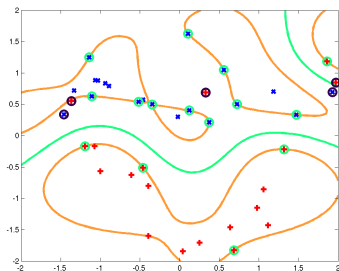
Chemin de régularisation

Suivi des $y_i \alpha_i$



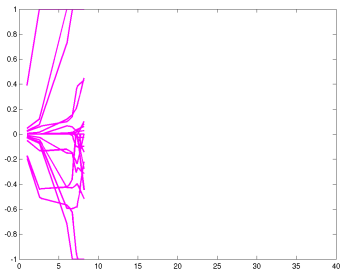
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



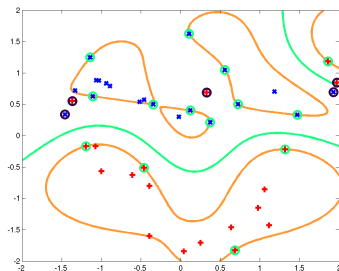
Chemin de régularisation

Suivi des $y_i \alpha_i$



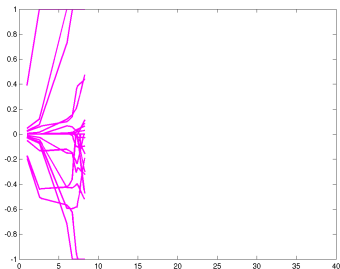
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



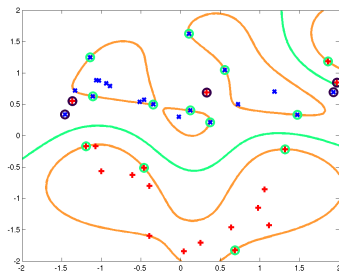
Chemin de régularisation

Suivi des $y_i \alpha_i$



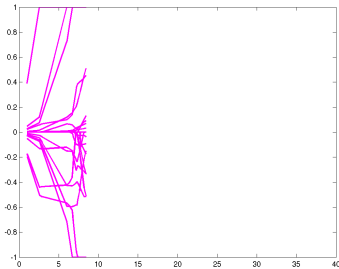
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



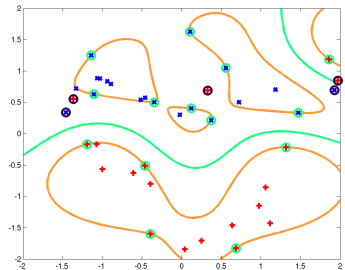
Chemin de régularisation

Suivi des $y_i \alpha_i$



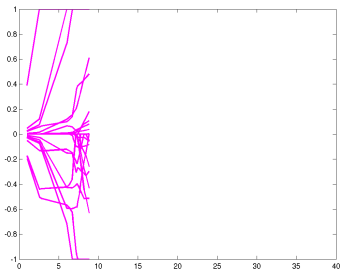
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



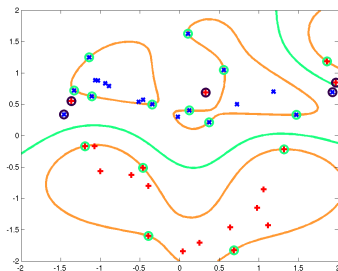
Chemin de régularisation

Suivi des $y_i \alpha_i$



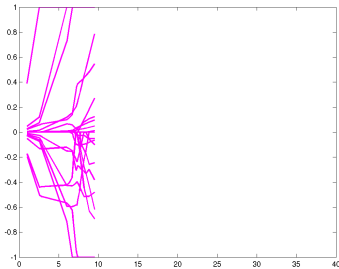
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



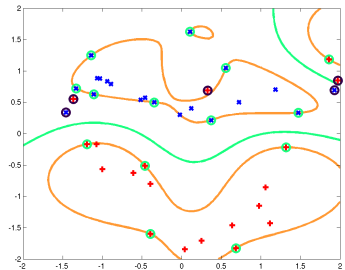
Chemin de régularisation

Suivi des $y_i \alpha_i$



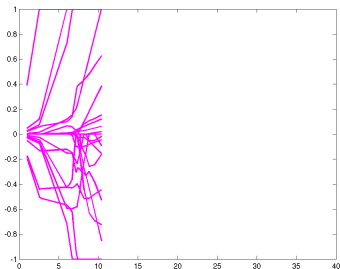
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



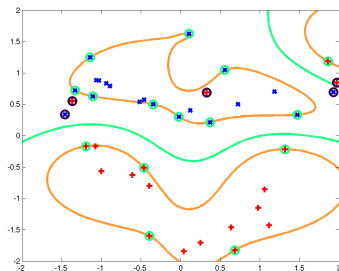
Chemin de régularisation

Suivi des $y_i \alpha_i$



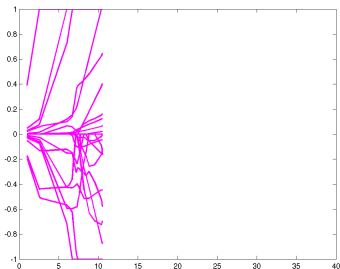
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



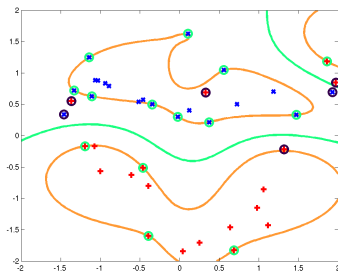
Chemin de régularisation

Suivi des $y_i \alpha_i$



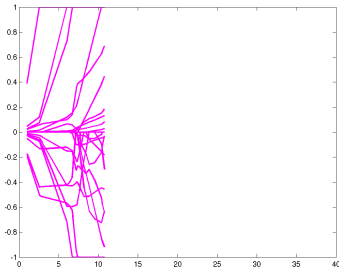
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



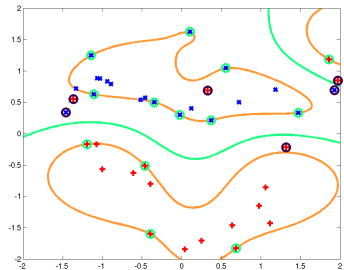
Chemin de régularisation

Suivi des $y_i \alpha_i$



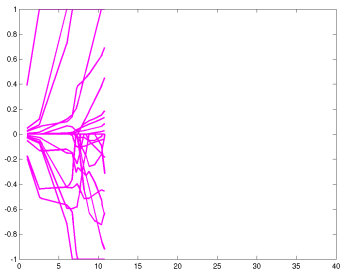
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



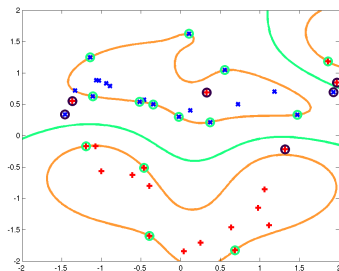
Chemin de régularisation

Suivi des $y_i \alpha_i$



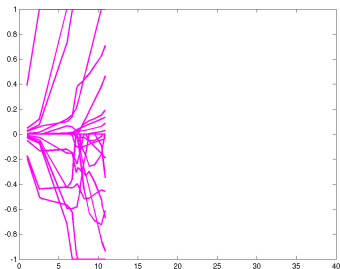
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



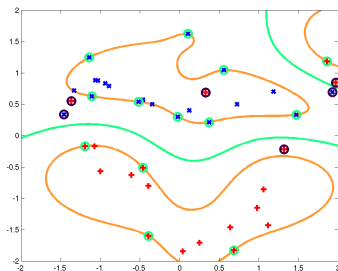
Chemin de régularisation

Suivi des $y_j \alpha_j$



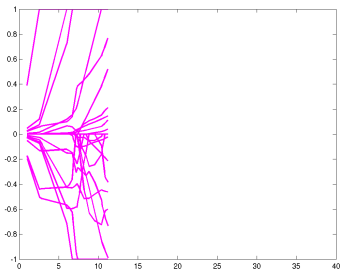
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



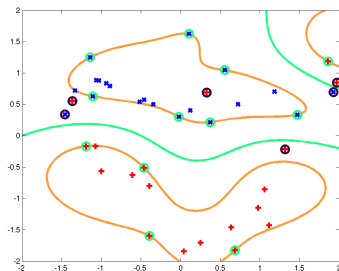
Chemin de régularisation

Suivi des $y_i \alpha_i$



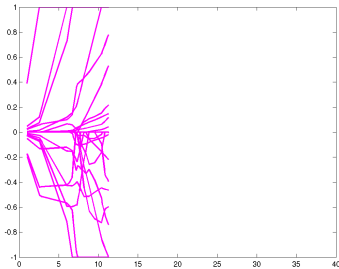
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



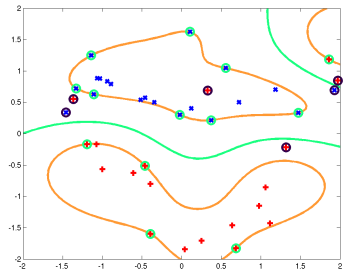
Chemin de régularisation

Suivi des $y_i \alpha_i$



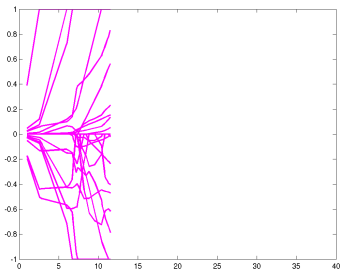
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



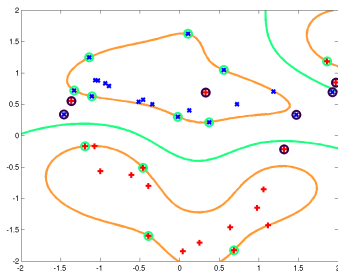
Chemin de régularisation

Suivi des $y_i \alpha_i$



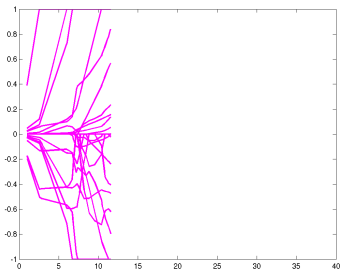
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



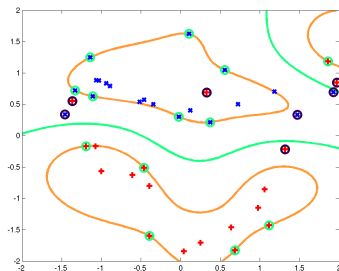
Chemin de régularisation

Suivi des $y_i \alpha_i$



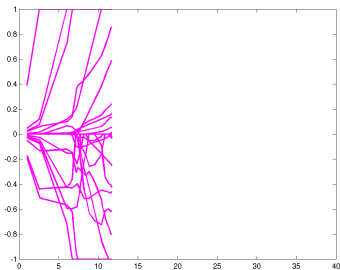
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



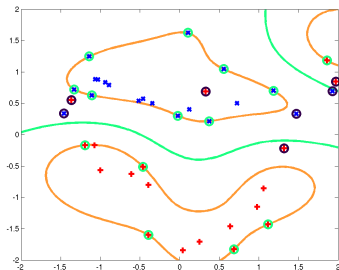
Chemin de régularisation

Suivi des $y_i \alpha_i$



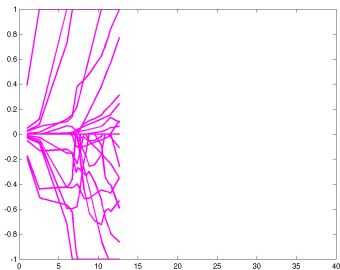
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



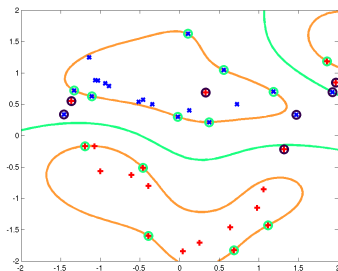
Chemin de régularisation

Suivi des $y_i \alpha_i$



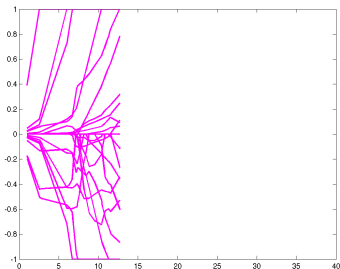
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



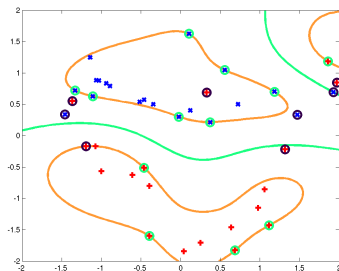
Chemin de régularisation

Suivi des $y_i \alpha_i$



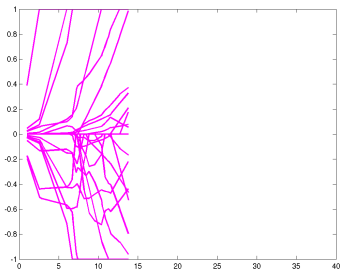
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



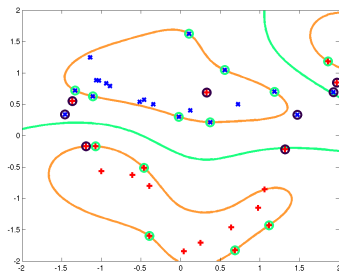
Chemin de régularisation

Suivi des $y_i \alpha_i$



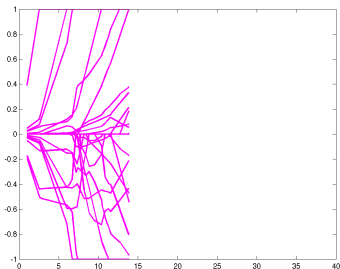
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



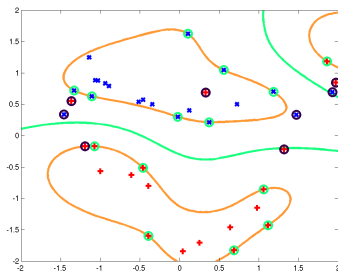
Chemin de régularisation

Suivi des $y_i \alpha_i$



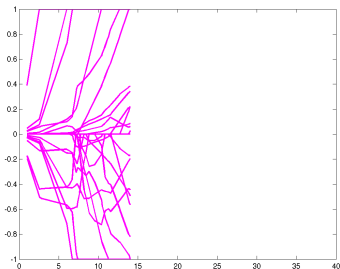
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



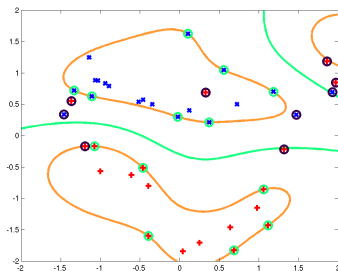
Chemin de régularisation

Suivi des $y_i \alpha_i$



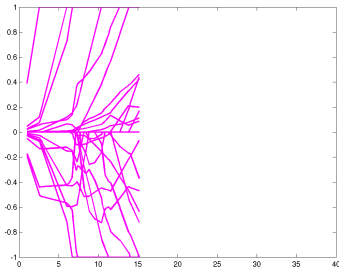
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



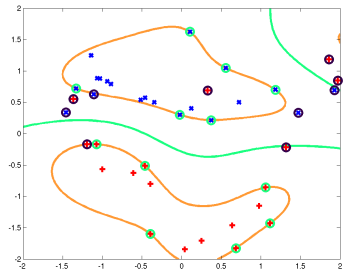
Chemin de régularisation

Suivi des $y_i \alpha_i$



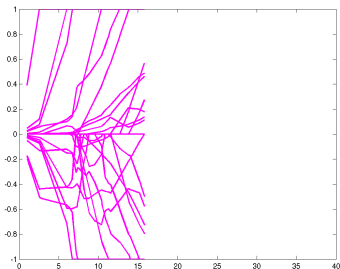
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



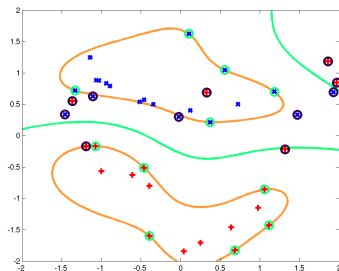
Chemin de régularisation

Suivi des $y_i \alpha_i$



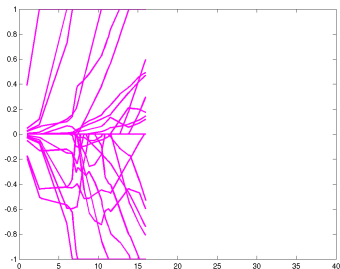
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



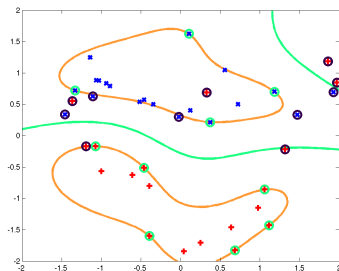
Chemin de régularisation

Suivi des $y_i \alpha_i$



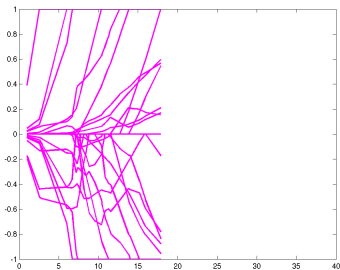
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



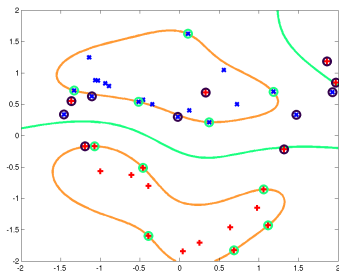
Chemin de régularisation

Suivi des $y_i \alpha_i$



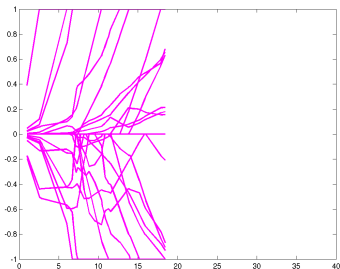
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



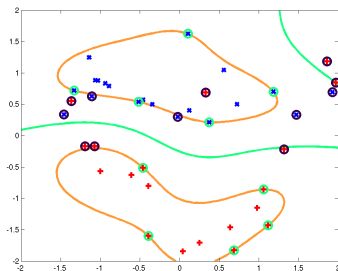
Chemin de régularisation

Suivi des $y_i \alpha_i$



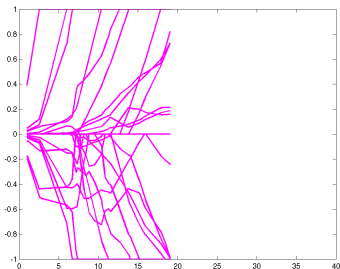
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



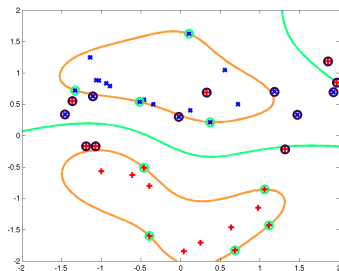
Chemin de régularisation

Suivi des $y_i \alpha_i$



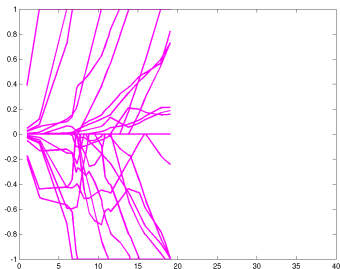
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



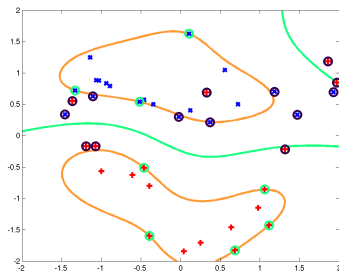
Chemin de régularisation

Suivi des $y_i \alpha_i$



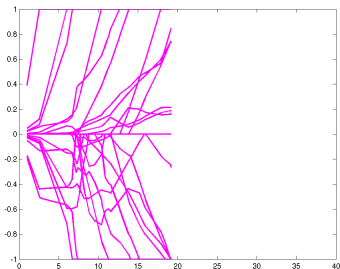
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



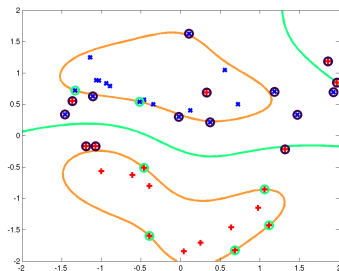
Chemin de régularisation

Suivi des $y_i \alpha_i$



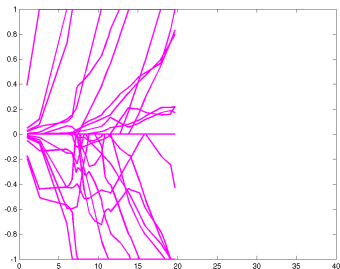
$$\nu = 1 (C \rightarrow \infty) \rightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



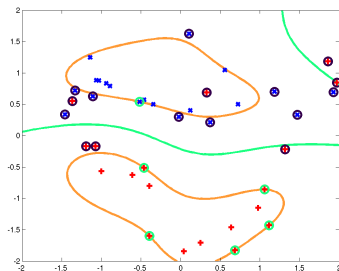
Chemin de régularisation

Suivi des $y_j \alpha_j$



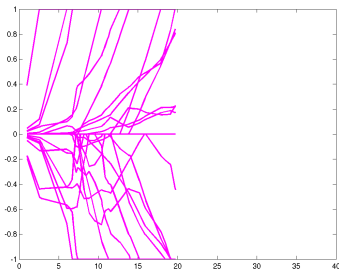
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



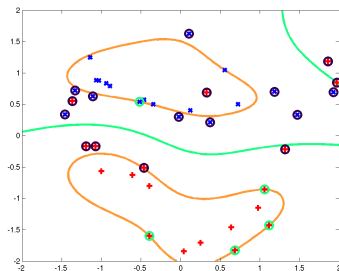
Chemin de régularisation

Suivi des $y_j \alpha_j$



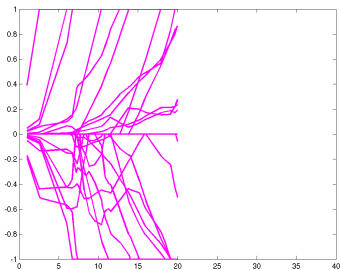
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



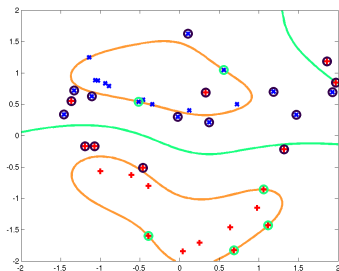
Chemin de régularisation

Suivi des $y_j \alpha_j$



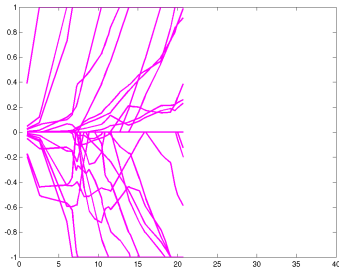
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



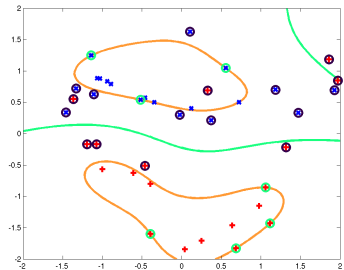
Chemin de régularisation

Suivi des $y_j \alpha_j$



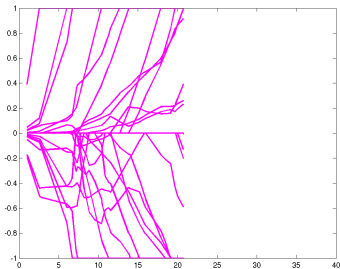
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



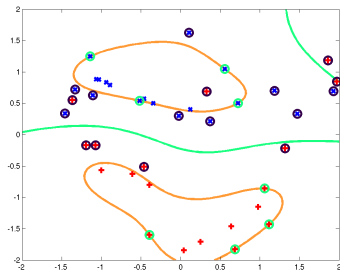
Chemin de régularisation

Suivi des $y_j \alpha_j$



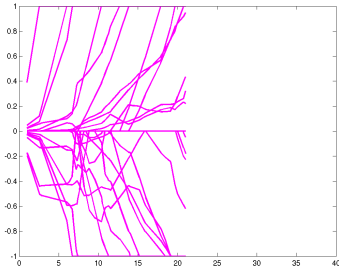
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



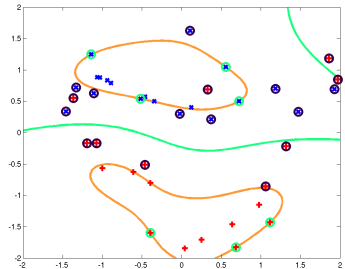
Chemin de régularisation

Suivi des $y_i \alpha_i$



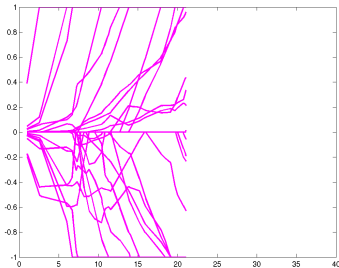
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



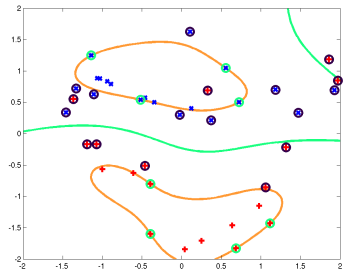
Chemin de régularisation

Suivi des $y_i \alpha_i$



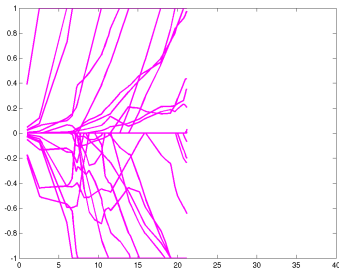
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



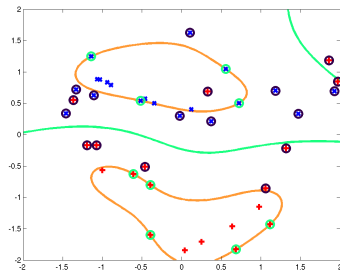
Chemin de régularisation

Suivi des $y_i \alpha_i$



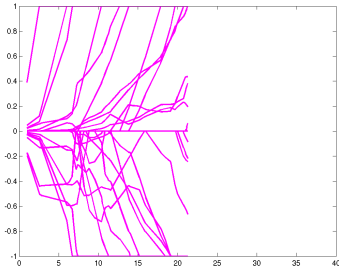
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



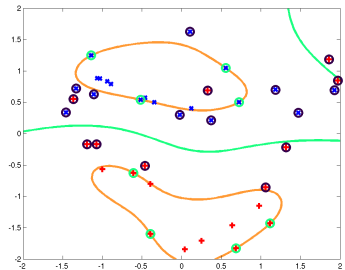
Chemin de régularisation

Suivi des $y_j \alpha_j$



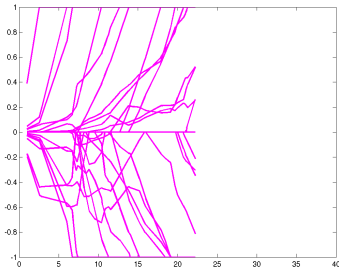
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



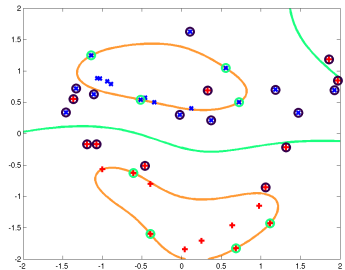
Chemin de régularisation

Suivi des $y_i \alpha_i$



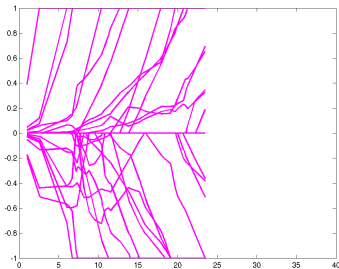
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



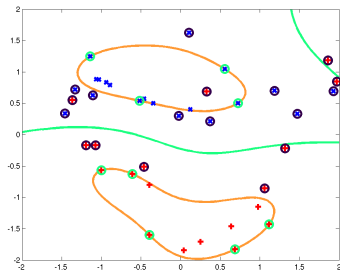
Chemin de régularisation

Suivi des $y_i \alpha_i$



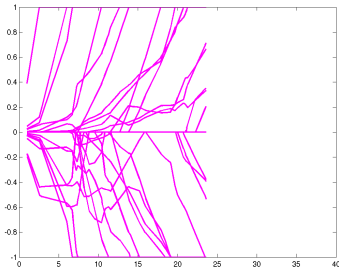
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



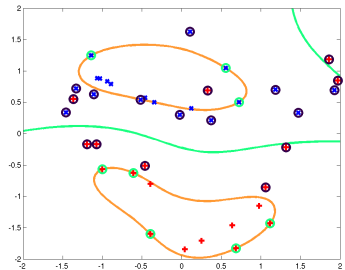
Chemin de régularisation

Suivi des $y_i \alpha_i$



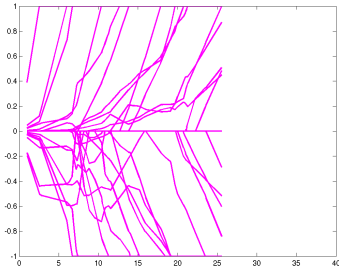
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



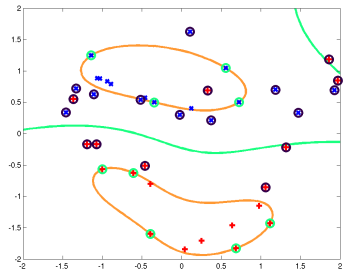
Chemin de régularisation

Suivi des $y_i \alpha_i$



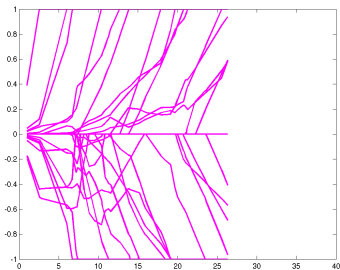
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



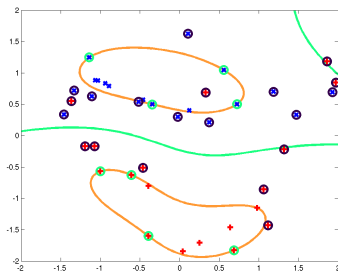
Chemin de régularisation

Suivi des $y_i \alpha_i$



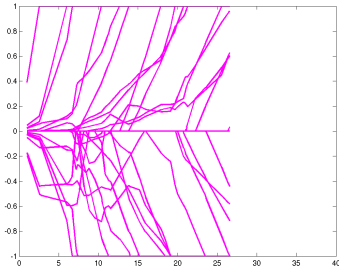
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



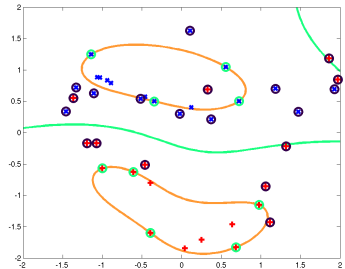
Chemin de régularisation

Suivi des $y_i \alpha_i$



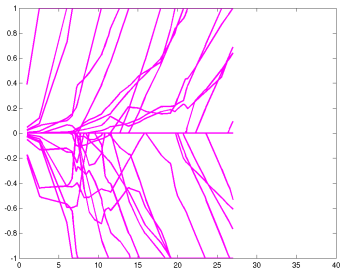
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



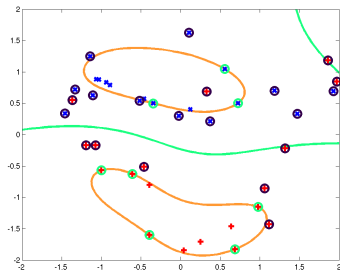
Chemin de régularisation

Suivi des $y_i \alpha_i$



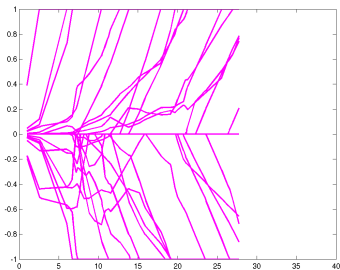
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



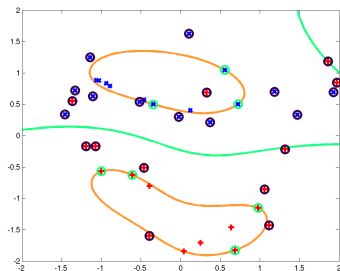
Chemin de régularisation

Suivi des $y_i \alpha_i$



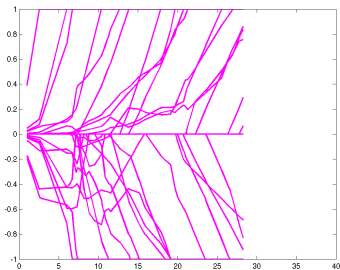
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



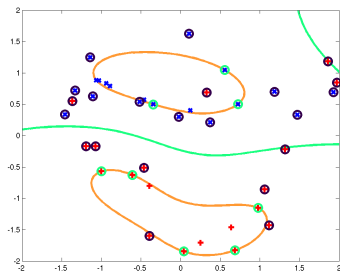
Chemin de régularisation

Suivi des $y_i \alpha_i$



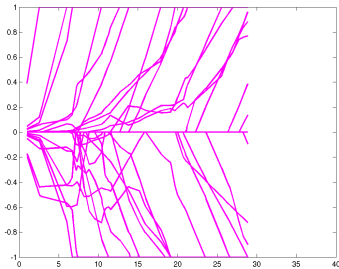
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



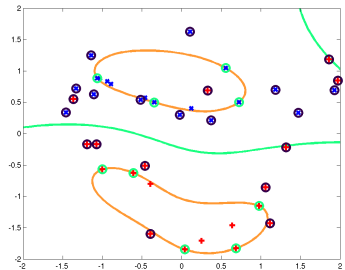
Chemin de régularisation

Suivi des $y_i \alpha_i$



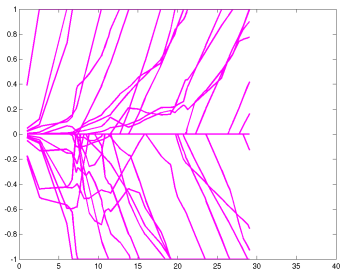
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



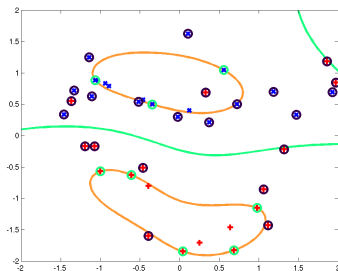
Chemin de régularisation

Suivi des $y_i \alpha_i$



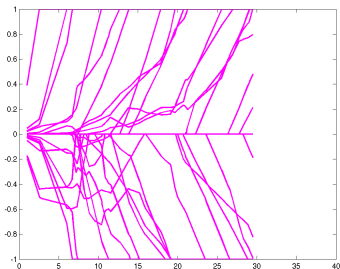
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



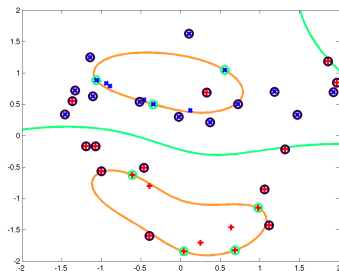
Chemin de régularisation

Suivi des $y_i \alpha_i$



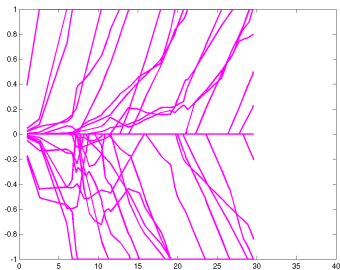
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



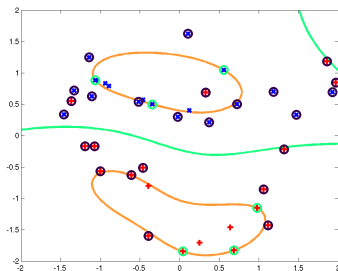
Chemin de régularisation

Suivi des $y_i \alpha_i$



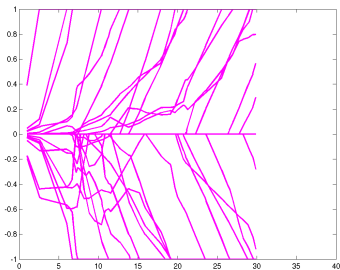
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



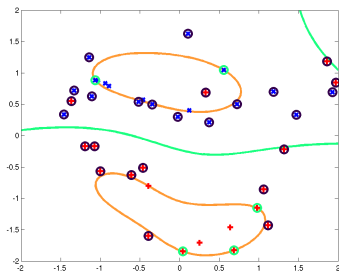
Chemin de régularisation

Suivi des $y_j \alpha_j$



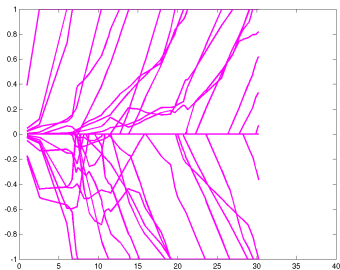
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



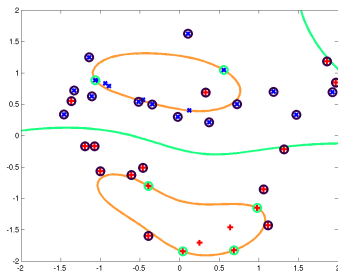
Chemin de régularisation

Suivi des $y_i \alpha_i$



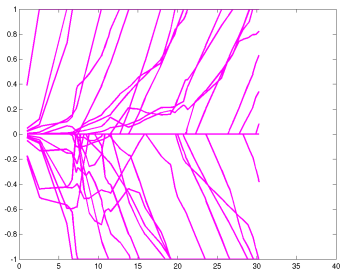
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



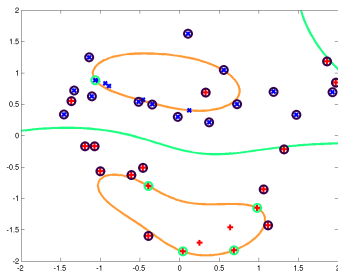
Chemin de régularisation

Suivi des $y_i \alpha_i$



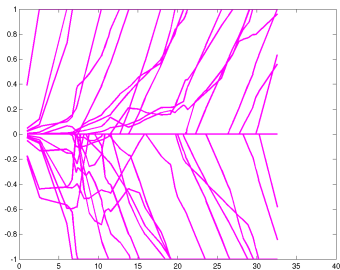
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



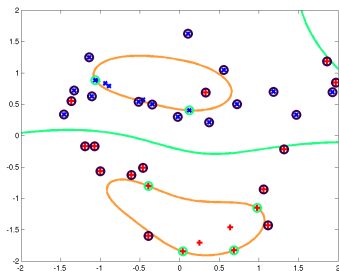
Chemin de régularisation

Suivi des $y_i \alpha_i$



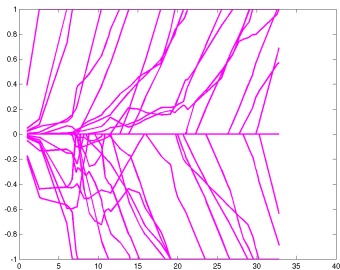
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



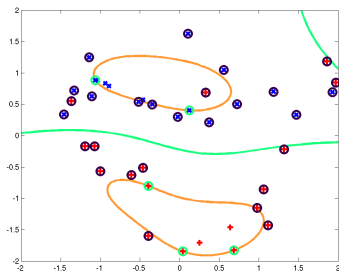
Chemin de régularisation

Suivi des $y_i \alpha_i$



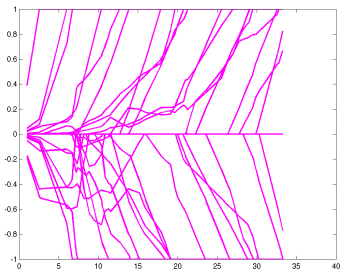
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



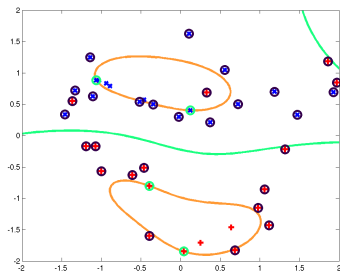
Chemin de régularisation

Suivi des $y_i \alpha_i$



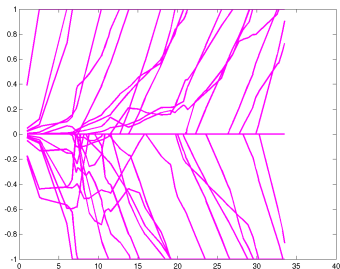
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



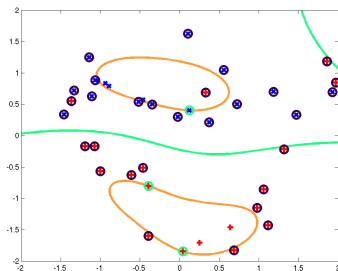
Chemin de régularisation

Suivi des $y_i \alpha_i$



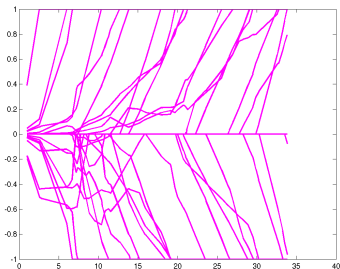
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



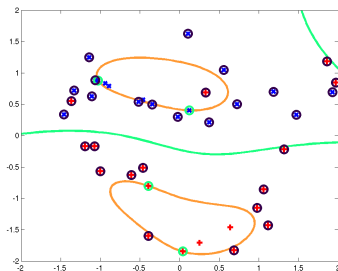
Chemin de régularisation

Suivi des $y_i \alpha_i$



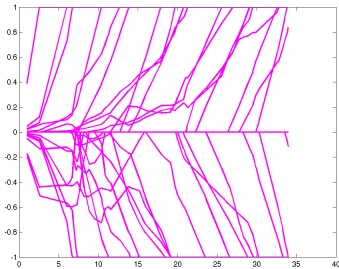
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



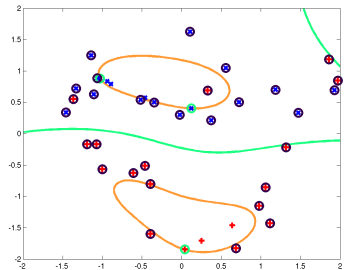
Chemin de régularisation

Suivi des $y_i \alpha_i$



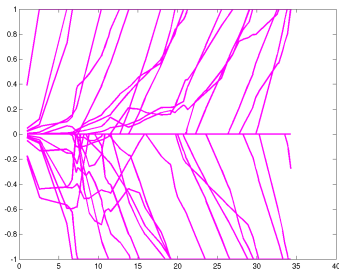
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



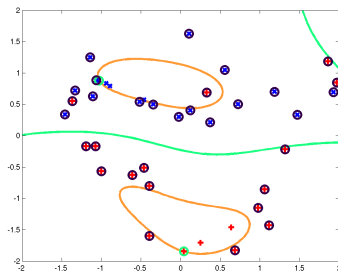
Chemin de régularisation

Suivi des $y_i \alpha_i$



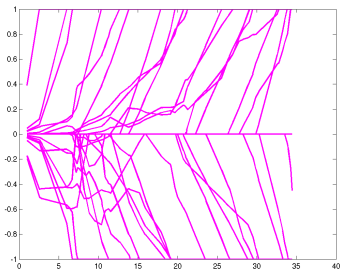
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



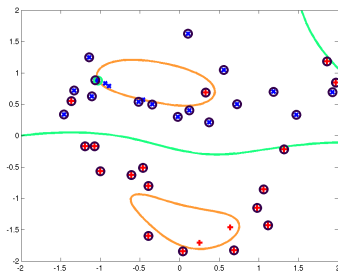
Chemin de régularisation

Suivi des $y_i \alpha_i$



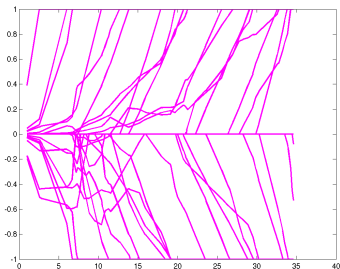
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



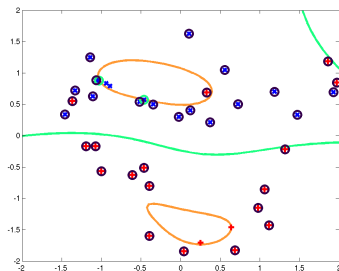
Chemin de régularisation

Suivi des $y_i \alpha_i$



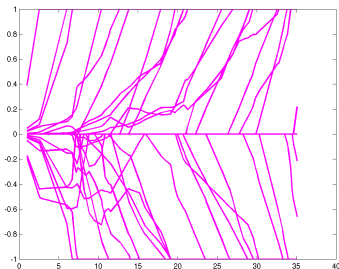
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



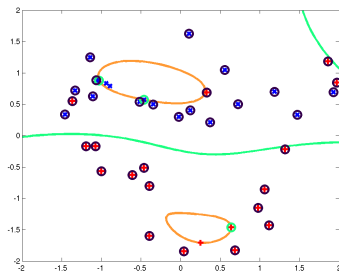
Chemin de régularisation

Suivi des $y_i \alpha_i$



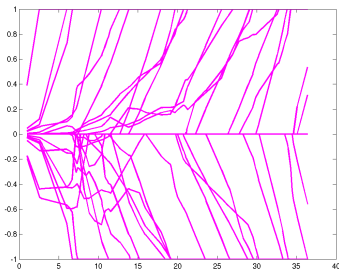
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



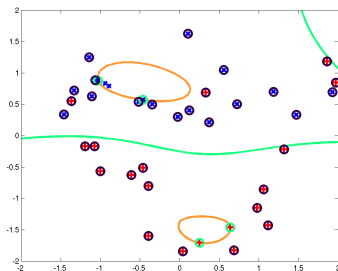
Chemin de régularisation

Suivi des $y_i \alpha_i$



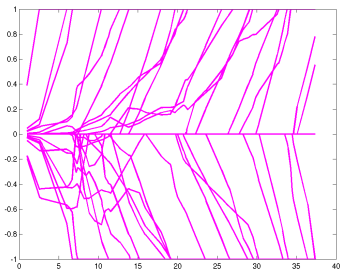
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



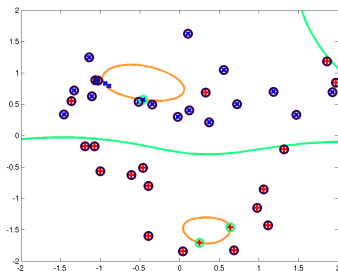
Chemin de régularisation

Suivi des $y_i \alpha_i$



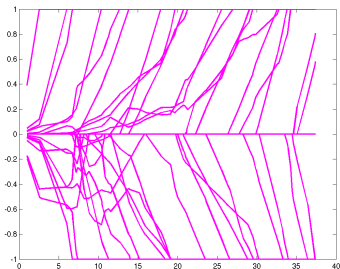
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



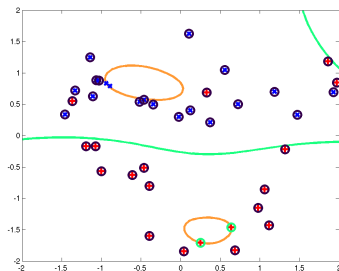
Chemin de régularisation

Suivi des $y_i \alpha_i$



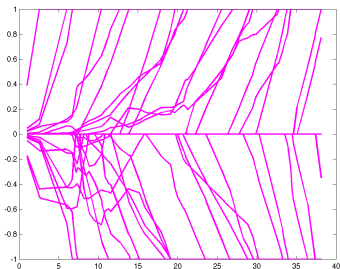
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



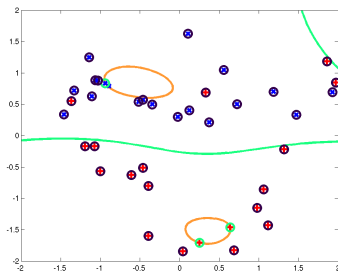
Chemin de régularisation

Suivi des $y_i \alpha_i$



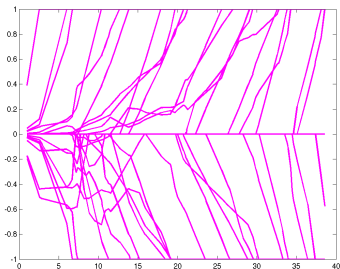
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



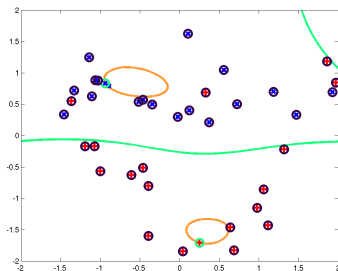
Chemin de régularisation

Suivi des $y_i \alpha_i$



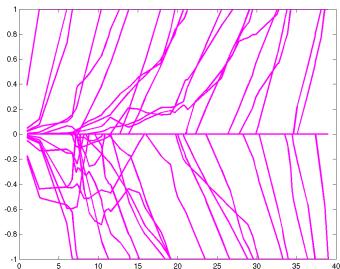
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



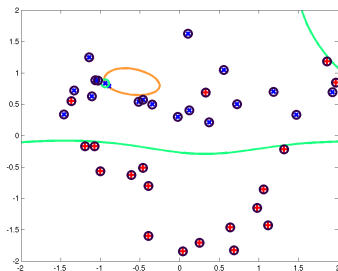
Chemin de régularisation

Suivi des $y_j \alpha_j$



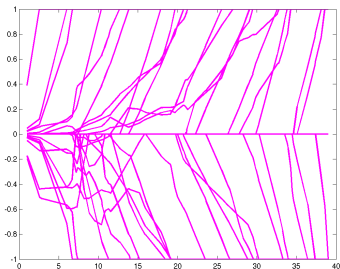
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



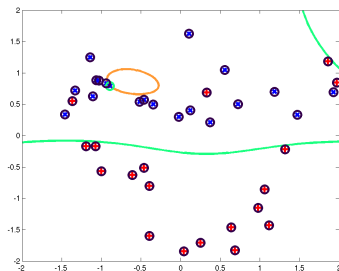
Chemin de régularisation

Suivi des $y_i \alpha_i$



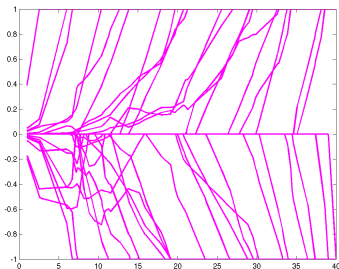
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



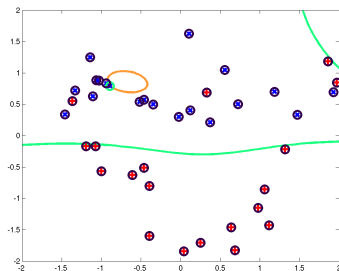
Chemin de régularisation

Suivi des $y_i \alpha_i$



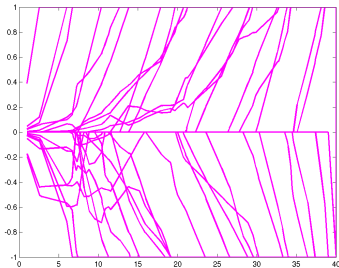
$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



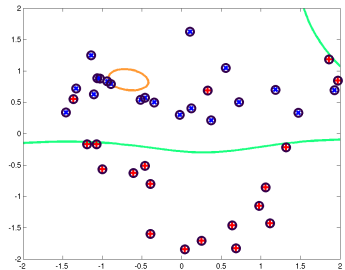
Chemin de régularisation

Suivi des $y_j \alpha_j$



$$\nu = 1 (C \rightarrow \infty) \longrightarrow \nu = n (C \rightarrow 0)$$

Suivi du chemin de régularisation



Chemin de régularisation des SVM

Principe [Hastie et al., 2004, Rosset & Zhu, 2006]

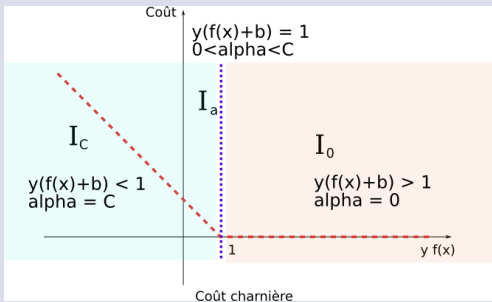
Vitesse de calcul

Utilité

Chemin de régularisation des SVM

Principe [Hastie et al., 2004, Rosset & Zhu, 2006]

- 1 la répartition des groupes I_a, I_C, I_0 est stable entre deux valeurs de ν



Vitesse de calcul

Utilité

Chemin de régularisation des SVM

Principe [Hastie et al., 2004, Rosset & Zhu, 2006]

Vitesse de calcul

- 1 chemin linéaire par morceaux : calcul d'un nombre fini de solutions
- 2 $\alpha^{t+1} = \alpha^t + (\nu^t - \nu^{t+1})\eta$

Utilité

Chemin de régularisation des SVM

Principe [Hastie et al., 2004, Rosset & Zhu, 2006]

Vitesse de calcul

Utilité

Déterminer la meilleure valeur de ν (ou de C)

Arrêt sur le chemin du ν -SVM

Pourquoi s'arrêter ?

Comment s'arrêter ?

Méthodes existante

Arrêt sur le chemin du ν -SVM

Pourquoi s'arrêter ?

- les dernières solutions ne sont pas parcimonieuses
- les dernières solutions sont presque équivalentes en classification

Comment s'arrêter ?

Méthodes existante

Arrêt sur le chemin du ν -SVM

Pourquoi s'arrêter ?

Comment s'arrêter ?

- évaluer la performance en généralisation à chaque étape
- évaluer le sur-apprentissage à chaque étape
- arrêter lorsque la solution est régulière et que l'erreur en généralisation est minimale

Méthodes existante

Arrêt sur le chemin du ν -SVM

Pourquoi s'arrêter ?

Comment s'arrêter ?

Méthodes existante

- calcul de bornes : calculs des C_p [Hastie et al., 2004], span bound [Chapelle et al., 2002], GACV [Lin et al., 2000]
- validation croisée, *leave-one-out*

Arrêt sur le chemin du ν -SVM

Pourquoi s'arrêter ?

Comment s'arrêter ?

Méthodes existante

- calcul de bornes : calculs des C_p [Hastie et al., 2004], span bound [Chapelle et al., 2002], GACV [Lin et al., 2000] **imprécis**
- validation croisée, *leave-one-out* **lent**

Evaluation par *leave-one-out*

Principe

Retirer un exemple de la base et apprendre avec les autres exemples. Faire la moyenne des erreurs commises sur chaque point retiré.

$$LOO1_{error} = \frac{1}{2n} \sum_i 1 - \text{sign}(\hat{y}_i y_i) \quad LOO2_{error} = \frac{1}{n} \sum_i \max(0, \rho_i - \hat{y}_i y_i)$$

- calcul des estimations de LOO à chaque étape du chemin
- les points de I_0 ne participent pas à la solution \Rightarrow l'erreur LOO est 0, pas de calcul
- on utilise la reprise à chaud du SimpleSVM pour calculer les solutions

Calcul très efficace pour les solutions parcimonieuses du chemin et toujours au moins aussi rapide qu'une évaluation externe

Evaluation par *leave-one-out*

Principe

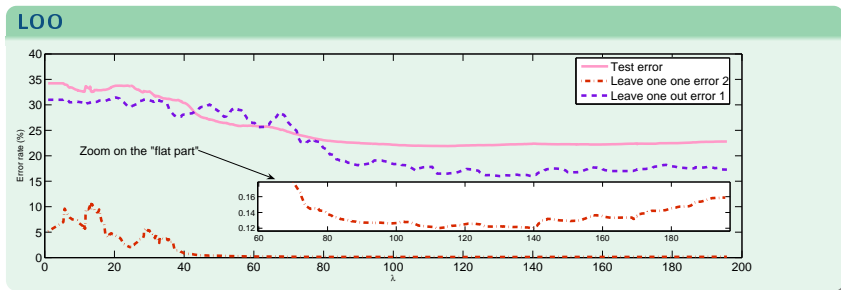
Retirer un exemple de la base et apprendre avec les autres exemples. Faire la moyenne des erreurs commises sur chaque point retiré.

$$LOO1_{error} = \frac{1}{2n} \sum_i 1 - \text{sign}(\hat{y}_i y_i) \quad LOO2_{error} = \frac{1}{n} \sum_i \max(0, \rho_i - \hat{y}_i y_i)$$

- calcul des estimations de LOO à chaque étape du chemin
- les points de l_0 ne participent pas à la solution \Rightarrow l'erreur LOO est 0, pas de calcul
- on utilise la reprise à chaud du SimpleSVM pour calculer les solutions

Calcul très efficace pour les solutions parcimonieuses du chemin et toujours au moins aussi rapide qu'une évaluation externe

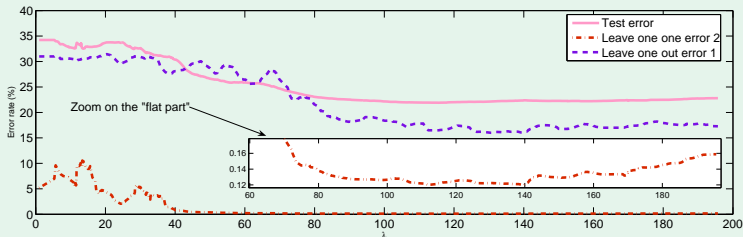
Illustration



Résultats sur le problème « mixture » [Hastie et al., 2004]

Illustration

LOO



Résultats sur le problème « mixture » [Hastie et al., 2004]



Gaëlle Loosli and Stéphane Canu.

Stopping on the ν -svm regularization using LOO.
Technical report, LITIS Rouen, 2006.

Conclusion

Détection de contexte

- 1 architecture de traitement de données
- 2 faisabilité par la détection de rupture

Apprentissage autonome

- 1 SimpleSVM : résolution efficace des SVM
- 2 gestion des très grandes bases de données via LASVM
- 3 chemin de régularisation parcouru dans le bon sens
- 4 réglage automatique de ν

Vers une méthode endurente

Objectifs

- fonctionnement en ligne
- adaptation en ligne des hyper-paramètres

Outils

- chemin de régularisation avec évaluation de la solution courante
- LASVM
- on sait rattraper le chemin de régularisation avec des étapes de SimpleSVM

A faire

- choisir dans quel sens parcourir le chemin
- gérer la capacité d'oubli
- il y a d'autres hyper-paramètres (le noyau) à traiter

Vers une méthode endurente

Objectifs

- fonctionnement en ligne
- adaptation en ligne des hyper-paramètres

Outils

- chemin de régularisation avec évaluation de la solution courante
- LASVM
- on sait rattraper le chemin de régularisation avec des étapes de SimpleSVM

A faire

- choisir dans quel sens parcourir le chemin
- gérer la capacité d'oubli
- il y a d'autres hyper-paramètres (le noyau) à traiter

Vers une méthode endurente

Objectifs

- fonctionnement en ligne
- adaptation en ligne des hyper-paramètres

Outils

- chemin de régularisation avec évaluation de la solution courante
- LASVM
- on sait rattraper le chemin de régularisation avec des étapes de SimpleSVM

A faire

- choisir dans quel sens parcourir le chemin
- gérer la capacité d'oubli
- il y a d'autres hyper-paramètres (le noyau) à traiter

Vers une méthode endurente

- 1: initialiser avec un premier ensemble de points (ν^* et solution associée)
- 2: **tant que** il y a des nouveaux points **faire**
- 3: mettre à jour la solution (via LASVM avec critère de sélection?)
- 4: **si** ν^* doit être modifié (augmenter ou diminuer ν ?) **alors**
- 5: mettre à jour ν^*
- 6: mettre à jour la solution
- 7: **fin si**
- 8: **fin tant que**

Merci



Basseville, M. & Nikiforov, I. V. (1993).

Detection of Abrupt Changes - Theory and Application.

Prentice-Hall.



Bordes, A., Ertekin, S., Weston, J., & Bottou, L. (2005).

Fast kernel classifiers with online and active learning.

Journal of Machine Learning Research, 6, 1579–1619.



Canu, S. & Smola, A. J. (2006).

Kernel methods and the exponential family.

Neurocomputing, 69(7-9), 714–720.



Chapelle, O., Vapnik, V., Bousquet, O., & Mukerjee, S. (2002).

Choosing multiple parameters for svm.

Machine Learning, 46(1-3), 131–159.



Desobry, F., Davy, M., & Doncarli, C. (2005).

An online kernel change detection algorithm.

IEEE Signal Processing, 53(5).



Hastie, T., Rosset, S., Tibshirani, R., & Zhu, J. (2004).

The entire regularization path for the support vector machine.

Journal of Machine Learning Research, 5, 1391–1415.



LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998).

Gradient-based learning applied to document recognition.

Proceedings of the IEEE, 86(11), 2278–2324.

<http://yann.lecun.com/exdb/mnist/>.

Lin, Y., Wahba, G., Hang, Z., & Lee, Y. (2000).

Statistical properties and adaptive tuning of Support Vector Machines.

Technical Report TR1022, Department of Statistics, Univ. Of Wisconsin, Madison WI.



Rosset, S. & Zhu, J. (2006).

Piecewise linear regularized solution paths.

Annals of Statistics.



Vishwanathan, S. V. N., Smola, A. J., & Murty, M. N. (2003).

SimpleSVM.

Proceedings of the Twentieth International Conference on Machine Learning.