

# PRMLの実装のすすめ

- PRML(Pattern Recognition and Machine Learning)とはビショップ著「パターン認識と機械学習」の本  
2006年の出版
- 最初はベイズ統計の本として知られ、現在は機械学習の教科書として認知されている。日本では専門書としては珍しく2万部売れている
- 統計数理の本として難易度が高いが、有名な手法は殆ど記述されているので多くの場所で勉強会が開かれている

# PRMLの実装のすすめ

- 自己紹介 @mabonakai0725  
20年以上データ分析  
統数研 持橋ゼミ(PRML§10担当)研究生
- PRMLとの出会い  
3年前 ベイズ統計勉強のためPRML独習  
2年前 筑波大(椿ゼミ)でPRML輪読  
2年前 持橋ゼミでMachine Learning(Kevin Murphy)PRMLの詳細版 輪読

# PRML感想1（雑感ですいません）

- 2006年の出版だが現在でも最先端と感じる。載ってないのは DeepLearning LDA ノンパラベイズぐらいと思う
- **基礎から応用まで完成度が高い**(基礎:我慢 応用:楽しく)
  - 驚くことにPRMLの後継図書を推奨できる人がいない
  - §2のガウス分布の分割、尤度、事後分布の公式は機械学習の基礎で、広範囲で利用するので習得が必要
  - §3、§4は機械学習の基本的な考え方が詰まってる
    - §3線形回帰 バリアンス(過学習) バイアス(劣学習) Lasso
    - §4線形識別 ラプラス近似 一般線形理論
  - §5以降は応用編 著名な実績あるモデルの解説

# PRML感想2(雑感ですいません)

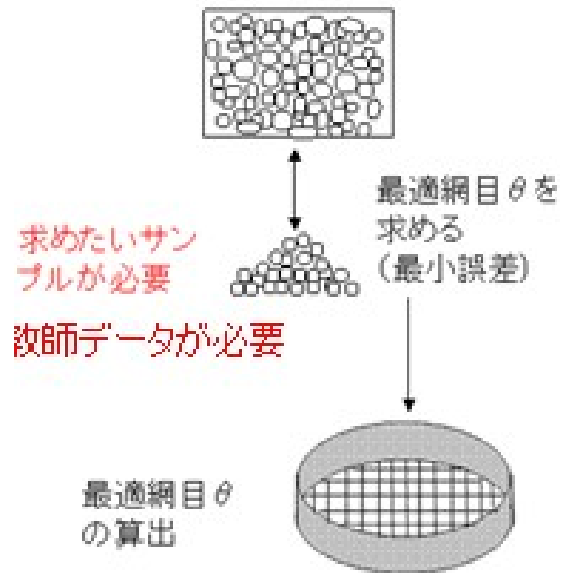
- ベイズ統計は未知な変数に事前分布を仮定し、EMやMCMCで事後分布を最大化して未知な変数(の挙動)を求める方式
- ビショップの熱意を感じる所  
パラメータの最適選択するARD(RMV) (各所に記述)  
変分とグラフィカルモデルと結合したERモデル(最長記述)
- ベイズ統計と頻度統計の相違

	ベイズ統計	頻度統計
モデル	生成モデル	識別モデル
データ	非教師付データ 教師付データ	教師付データ
手法	K-MEANS MCMC EM 隠れマルコフ	SVM ロジット回帰 線形回帰 ニューロ

# 何故ベイズ統計が機械学習か

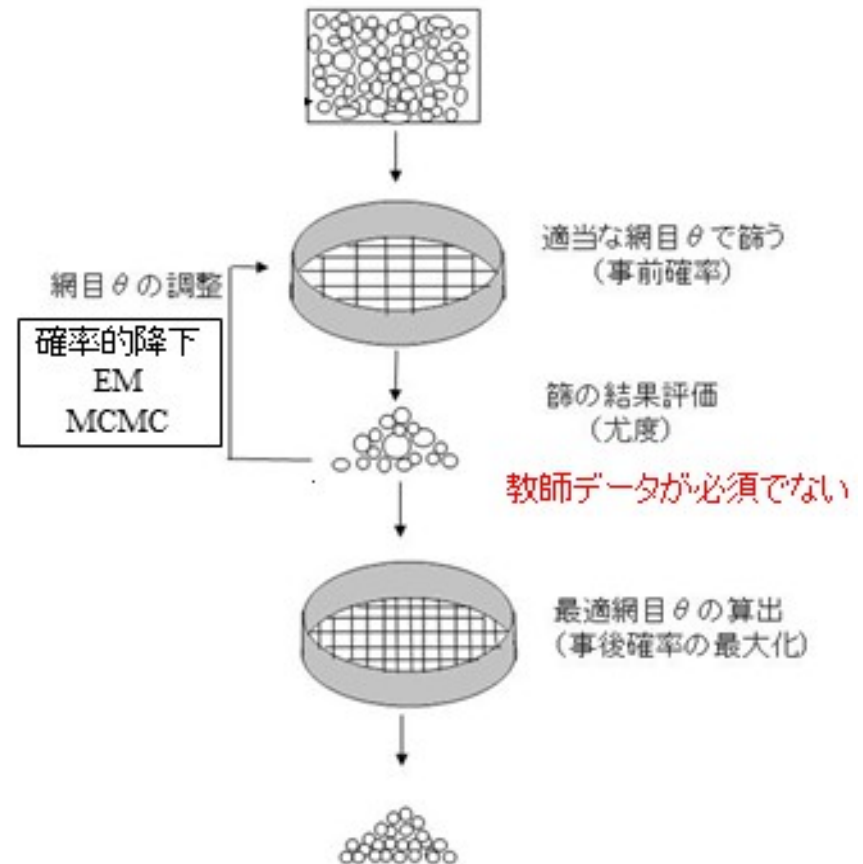
## 識別モデル

$\max p(\mathbf{X}|\theta)$   
 $\theta$ : パラメータ  
 $\mathbf{X}$ : データ



## 生成モデル

$\max p(\theta|\mathbf{X}) \propto p(\mathbf{X}|\theta)p(\theta)$   
 $\theta$ : パラメータ、隠変数  
 $\mathbf{X}$ : データ



データに合わせて学習

# PRMLの実装 1

- PRMLの式を理解するためにC言語で実装した
  - 殆ど文献にある式通りに実装できるので
    - 式の省略 誤植が無い証拠
  - 一見難しい式でも意外に簡単に実装できる
  - 理解せずとも式をプログラムするだけで動く場合がある
    - ガウス回帰、ガウス識別
  - 難解な式もプログラム(アルゴリズム)にすると理解し易い
  - 文献通り動くと理解した気になる
    - 隠れマルコフモデル(理論は難しいが、プログラムは簡単)

# PRMLの実装 2

- ネット上にPRML(x.xx)と検索すると解説がヒットする
  - 式の展開を理解するには非常に助けとなる
- 難しいモデルの実装はネット上にあるHPを参照した
  - サイボーズの中谷さん、京大の近藤誠一さん  
[http://d.hatena.ne.jp/n\\_shuyo/20121004/prml](http://d.hatena.ne.jp/n_shuyo/20121004/prml)
  - 優秀なプログラムがネット上に多数存在する
- 重要だが実装できないモデルがあった
  - EP法、ノンパラベイズ

# 実装例1 ロジット回帰(SGD版)

download <http://www1.m.jcnnet.jp/mabonki/download.htm>

一般にロジット回帰は多次元ニュートン法で解くが、SGD法だと5行でできる。

繰返計算

```
for(ii=0;ii<iter;ii++) {  
    for(i=0;i<n;i++) {  
        sum=w[0];  
        for(j=0;j<m-1;j++) {  
            sum += w[j+1]*x[i][j];  
        }  
        #ifdef REG  
        ww[0] = w[0] + u * (y[i] - sum);  
        for(j=0;j<m-1;j++) {  
            ww[j+1] = w[j+1] + u*(y[i]-sum)*x[i][j];  
        }  
        #else  
        g=(1.0/(1+exp(-sum)));  
        ww[0] = w[0] + u * (y[i] - g);  
        for(j=0;j<m-1;j++) {  
            ww[j+1] = w[j+1] + u * (y[i] - g) * x[i][j];  
        }  
        #endif  
        /* 入れ替え */  
        for(j=0;j<m;j++) {  
            w[j]=ww[j];  
        }  
    }  
}
```

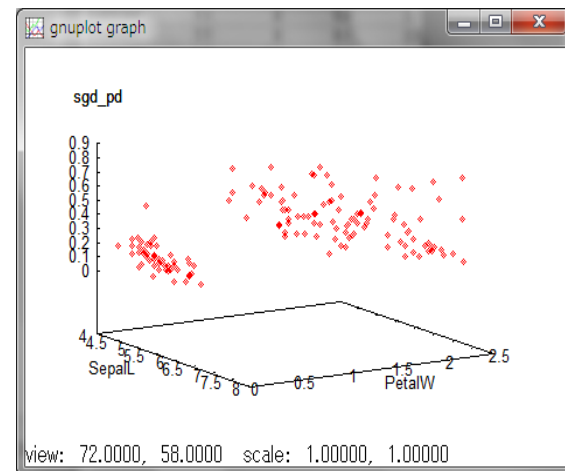
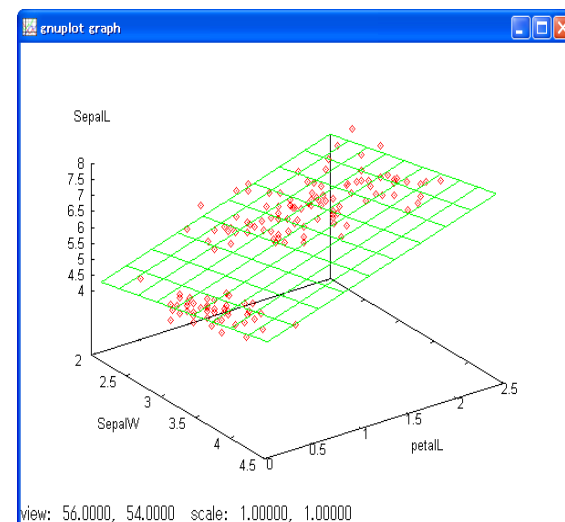
重回帰計算部分

ロジット回帰計算部分

重みの更新

教師区分

説明変数

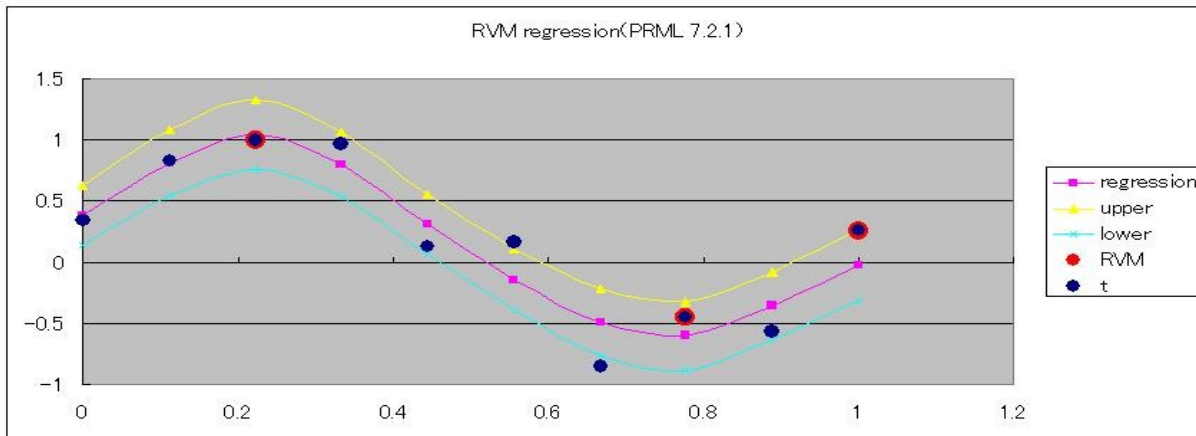




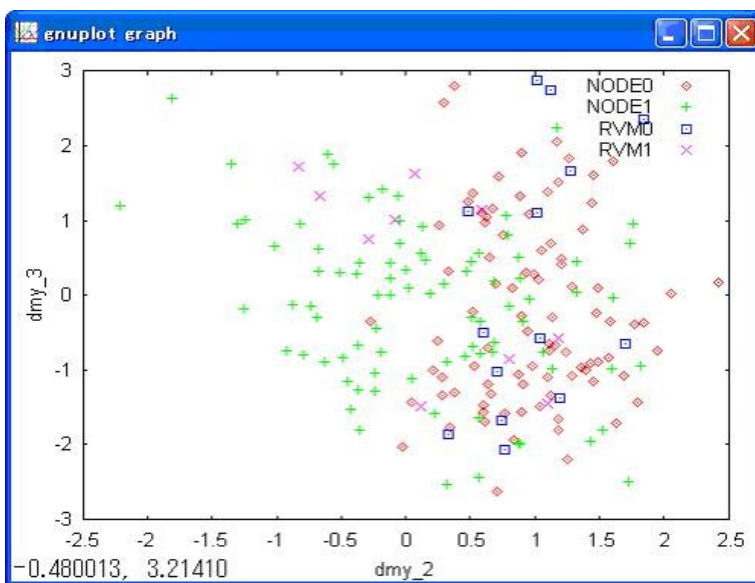
# 実装例2 ARD(関連度自動決定)RVM

download <http://www1.m.jcnnet.jp/mabonki/download.htm>

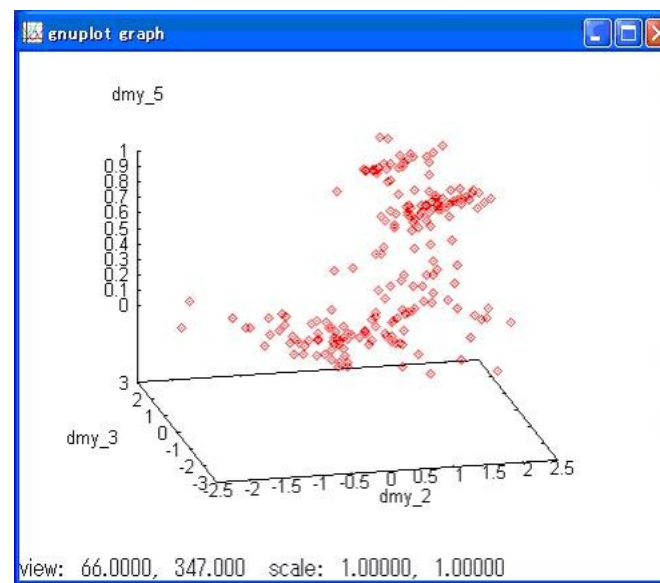
RVM回帰 出来るだけ少ない点で回帰する 下記例では◎の3点のみで回帰



RVM識別 □とXのみで識別



RVM識別の結果 Z軸は確率値



	章		節	実装したモデル	容易度
1	序論	1.2	曲線フィッティング	リッジ回帰モデル	◎
2	確率分布	2.5	最近近傍法	k-NN区分モデル	◎
3	線形回帰モデル	3.1.3	逐次学習	SGDIによるロジット回帰モデル	◎
		3.1.4	正則化最小二乗法	Lasso回帰モデル	2次最適化
4	線形識別モデル	4.3.2	ロジステック回帰	ロジステック回帰モデル	◎
		4.3.4	多クラスロジット回帰	多クラスロジット回帰モデル	○
5	ニューラルネットワーク	5.3	誤差逆伝播	ニューロモデル	○
		5.5.6	たたみ込みニューラルネット	Convolution DeepLearning	×
6	カーネル法	6.4	ガウス回帰	ガウス回帰モデル	◎
		6.4.5	ガウス過程による識別	ガウス識別モデル	◎
7	疎な解を持つカーネルマシン	7.1.1	重なり合うクラス分類	SVM-SMO	SMO
		7.1.3	多クラスSVM	1クラスSVM	2次最適化
		7.1.4	回帰のためのSVM	SVM回帰モデル	2次最適化
		7.2.1	回帰問題に対するRVM	RVM回帰モデル	○
		7.2.3	分類問題に対するRVM	RVM識別モデル	○
8	グラフィカルモデル	8.1.4	線形ガウスモデル	ガウシアン・グラフィカルモデル	△
		8.4.1	連鎖による推論(有向)	ベイジアンネットワークモデル	×
			連鎖による推論(無向)	MRF	×
8.4.8	グラフ構造学習	スコアベース構造学習	×		
9	混合モデルとEM	9.1	k-meansクラスタリング	k-means法	○
		9.2.2	混合ガウス分布のEM	EMモデル	◎
10	近似推論法	10.2	例:変分混合ガウス分布	変分法クラス分類モデル	○
11	サンプリング	11.2.2	Metero-Hestingsアルゴリズム	酔歩MCMC	○
		11.3	ギブスサンプリング	ギブスサンプリング	◎
12	連続潜在変数	12.1	主成分分析	PCAモデル	○
		12.2.4	因子分析	因子分析モデル	○
		12.4.1	独立成分分析	独立成分分析モデル	○
13	系列データ	13.2	隠れマルコフモデル	HMMモデル	○
		13.3	線形動的システム	カルマンフィルター	×
		13.3.4	粒子フィルター	粒子フィルター	△
14	モデルの結合	14.3	ブースティング	Adaブーストモデル	○
		14.4	木構造モデル	C4.5判別ツリー	○

# PRMLの次へ

- 実装できてもPRMLにあるデータはToy Data
  - 実装できても実用的なモデルかは疑問
  - カーネルモデル(SVM等)はデータが多いと動かない
- 現実のデータは複雑でツールを組合せる必要がある
  - 大規模で多次元だがデータが疎 → 次元縮約、SVD
  - 多峰性の分布 → MCMC
  - 潜在変数を仮定 → 変分近似 ラプラス近似 EM
  - 変数間が非線形 → カーネル法
- PRMLの知識は機械学習分野で共通話題となるが、どの程度理解しているかでテーマや指導者の選択が決まってしまう恐れがある。