2009/01/28　パターン認識特論演習問題　学籍番号＿＿＿＿＿＿＿＿＿＿氏名＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

1. Exclusive OR のパターンを学習することができる TLU ネットワークを図示し, その構造の必然性を説明しなさい。

2. 下記の表のようなテスト-目標属性間の関係があるとき，エントロピーゲインに基づいて決定木を求めなさい.

T1	T2	T3	T4	目標属性
1	0	1	1	1
1	1	1	1	1
1	1	1	0	0
1	0	1	0	0
1	1	0	1	0
1	0	0	1	0
1	1	0	1	0
0	1	0	1	1
0	0	1	1	1
0	1	0	1	1
0	0	0	1	1
0	0	1	0	0
0	1	0	0	0
0	0	1	0	0

3. d 次元ベクトル $\boldsymbol{f}_i \ (i = 1, \cdots, M)$ が与えられている．その固有ベクトルを求める方法について考える．$F = \left[\boldsymbol{f}_i \quad \cdots \quad \boldsymbol{f}_M \right]$ とすると，共分散行列を M 倍した行列は $\sum_{i=1}^{M} \boldsymbol{f}_i \boldsymbol{f}_i^T = FF^T$ と表せる．この行列の固有ベクトルを FF^T ではなく $F^T F$ の固有ベクトルから求める方法と，その利点について述べなさい.

4. $f(x)$ の Fourier 変換を $F(\omega)=\int_{-\infty}^{\infty} f(x)e^{-jwx}dx$ とする．このとき，$f(x)$ が平行移動したパターン $g(x)=f(x-\mu)$ の Fourier 変換を $F(\omega)$ を用いた式で表しなさい．また，μ に対して不変な特徴を抽出する方法の一つとして，自己相関関数 $R(\tau)=\int_{-\infty}^{\infty} f(x-\mu+\tau)f(x-\mu)dx$ が考えられる．$R(\tau)$ の Fourier 変換が $\| F(\omega) \|^2$ となることを示しなさい．また，この不変特徴の問題点について説明しなさい．

5. パターン $\mathbf{x}$ を正規直交基底 $\mathbf{u}_i\,(i=1,\cdots,n)$ が張る部分空間に正射影する行列 P を $U=\begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{pmatrix}$ という行列を用いて表現しなさい．

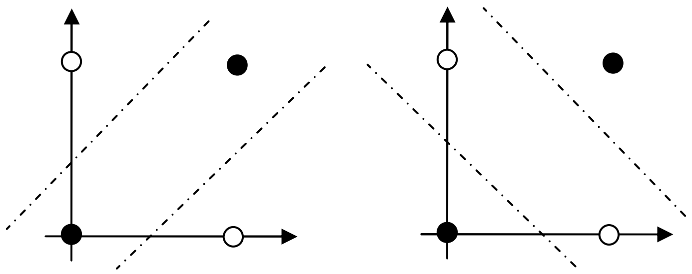
6. 5 において、C 個の部分空間が存在する場合，各部分空間への射影行列 P_i（$i=1,\cdots,C$）を用いて，「$\boldsymbol{x}^T P_i \boldsymbol{x}$ を最大化する i を求める問題」と，「$\| \boldsymbol{x}-P_i\boldsymbol{x} \|$ を最小化する i を求める問題」は本質的に同じであることを示しなさい．

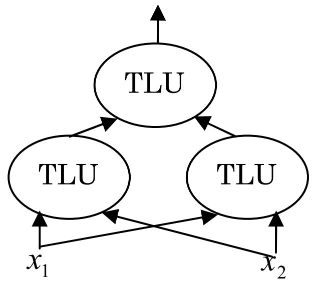
7. SVM の学習時に必要となる計算量・メモリ量がデータ数およびデータの次元数とどのような関係にあるかについて述べなさい．

8. 共分散行列 Σ の固有値と固有ベクトルがそれぞれ λ_i と $\boldsymbol{\varphi}_i\,(i=1\cdots n)$ であるとき，共分散行列 Σ を λ_i と $\boldsymbol{\varphi}_i$ を用いて表しなさい．

2009/01/28　パターン認識特論演習問題　学籍番号＿＿解答例＿＿氏名＿＿＿＿＿＿＿＿＿＿＿＿

1. Exclusive OR のパターンを学習することができる TLU ネットワークを図示し，その構造の必然性を説明しなさい。

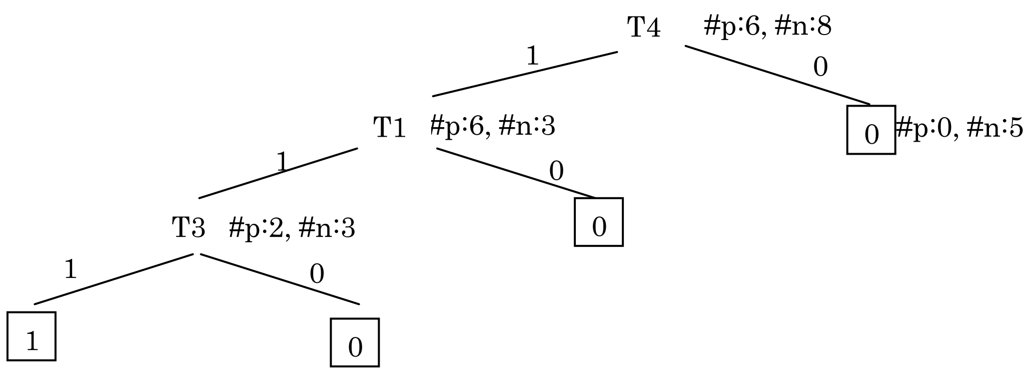




TLU は線形識別面の学習を行う機構である。線形識別面で Exclusive OR のパターンを表現しようとすると、少なくとも 2 つの TLU が必要になる。どちらのケースでも、線形識別面によって区切られるどちらの半空間にパターンが属するのかを判定する TLU と、それらを総合して、出力として 1 を出すか否かを決定する TLU が存在する必要があるため、上図の構造のネットワークになる。但し、各 TLU の-1 に固定された入力は省略している。

2. 下記の表のようなテスト-目標属性間の関係があるとき，エントロピーゲインに基づいて決定木を求めなさい.

T1	T2	T3	T4	目標属性
1	0	1	1	1
1	1	1	1	1
1	1	1	0	0
1	0	1	0	0
1	1	0	1	0
1	0	0	1	0
1	1	0	1	0
0	1	0	1	1
0	0	1	1	1
0	1	0	1	1
0	0	0	1	1
0	0	1	0	0
0	1	0	0	0
0	0	1	0	0



3. d 次元ベクトル $\boldsymbol{f}_i\ (i=1,\cdots,M)$ が与えられている．その固有ベクトルを求める方法について考える．$F=\left[\boldsymbol{f}_i \quad \cdots \quad \boldsymbol{f}_M\right]$ とすると，共分散行列を M 倍した行列は $\sum_{i=1}^{M}\boldsymbol{f}_i\boldsymbol{f}_i^T = FF^T$ と表せる．この行列の固有ベクトルを FF^T ではなく F^TF の固有ベクトルから求める方法と，その利点について述べなさい.

F^TF の固有ベクトルを φ，固有値を λ とすると，$F^TF\varphi=\lambda\varphi$ が成立する．この式の左から F をかけると，$FF^TF\varphi=\lambda F\varphi$ が得られる．$FF^T(F\varphi)=\lambda(F\varphi)$ と見なすと，FF^T の固有ベクトルは，$F\varphi$ と表せる．この方法のメリットは，$d \gg M$ の場合，FF^T が $d\times d$ であるのに対し F^TF は $M\times M$ と小さくなることである.

4. $f(x)$ の Fourier 変換を $F(\omega)=\int_{-\infty}^{\infty} f(x)e^{-jwx}dx$ とする．このとき，$f(x)$ が平行移動したパターン $g(x)=f(x-\mu)$ の Fourier 変換を $F(\omega)$ を用いた式で表しなさい．また，μ に対して不変な特徴を抽出する方法の一つとして，自己相関関数 $R(\tau)=\int_{-\infty}^{\infty} f(x-\mu+\tau)f(x-\mu)dx$ が考えられる．$R(\tau)$ の Fourier 変換が $\| F(\omega) \|^2$ となることを示しなさい．また，この不変特徴の問題点について説明しなさい．

$$G(\omega)=\int_{-\infty}^{\infty} f(x-\mu)e^{-jwx}dx=\int_{-\infty}^{\infty} f(X)e^{-jw(X+\mu)}dX=e^{-jw\mu}\int_{-\infty}^{\infty} f(X)e^{-jwX}dX=e^{-jw\mu}F(\omega)$$

$$\int_{-\infty}^{\infty} R(\tau)e^{-j\omega\tau}d\tau=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(x-\mu+\tau)f(x-\mu)dx\, e^{-j\omega\tau}d\tau=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(X+\tau)f(X)dX\, e^{-j\omega\tau}d\tau$$

変数変換を行い，次の結果を得る．

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f(T)f(X)dXe^{-j\omega(T-X)}dT=\int_{-\infty}^{\infty} f(T)e^{-j\omega T}dT\int_{-\infty}^{\infty} f(X)e^{j\omega X}dX=F(\omega)F^*(\omega)=\| F(\omega) \|^2$$

このようにして得られた不変特徴は，各周波数成分間の位相情報まで消されているため，識別に利用できる特徴が少なくなってしまうという問題点がある．

5. パターン $\mathbf{x}$ を正規直交基底 $\mathbf{u}_i\,(i=1,\cdots,n)$ が張る部分空間に正射影する行列 P を $U=\begin{pmatrix}\mathbf{u}_1 & \cdots & \mathbf{u}_n\end{pmatrix}$ という行列を用いて表現しなさい．

$$P=UU^T$$

6. 5 において、C 個の部分空間が存在する場合，各部分空間への射影行列 P_i（$i=1,\cdots,C$）を用いて，「$\boldsymbol{x}^T P_i \boldsymbol{x}$ を最大化する i を求める問題」と，「$\| \boldsymbol{x}-P_i\boldsymbol{x} \|$ を最小化する i を求める問題」は本質的に同じであることを示しなさい．

$\| \boldsymbol{x}-P_i\boldsymbol{x} \|$ を最小化する問題は，$\| \boldsymbol{x}-P_i\boldsymbol{x} \|^2$ を最小化することと同値である．これを計算すると，$\| \boldsymbol{x}-P_i\boldsymbol{x} \|^2=(x-P_ix)^T(x-P_ix)=\| x \|^2-2x^TP_ix+x^TP_i^TP_ix$ となる．$P_i=U_iU_i^T$ から $P_i^TP_i=P_i$ となるので，結局 $\| \boldsymbol{x}-P_i\boldsymbol{x} \|^2=\| x \|^2-x^TP_ix$ となる．このうち，x は i に依存せず最小化とは無関係であるので，$\| x \|^2$ を無視することができ，$-x^TP_ix$ を最小化すればよいことが分かる．これは，x^TP_ix を最大化する問題となっている．

7. SVM の学習時に必要となる計算量･メモリ量がデータ数およびデータの次元数とどのような関係にあるかについて述べなさい．

SVM では各データ間の内積によって定義されるマトリクスを用いた 2 次形式を含む目的関数の最適化を行う．このため，データ数 n に対してマトリクスのサイズは n^2 となる．また，この最適化問題を解くための計算量は $O(n^2)$ となる．データの次元数 d は実質的にカーネル関数を計算する際にしか問題にならず，この計算量は $O(dn^2)$ となる．したがって，次元数が大きい場合でもデータ数が多くなければ，学習を行うことができる．

8. 共分散行列 Σ の固有値と固有ベクトルがそれぞれ λ_i と $\boldsymbol{\varphi}_i\ (i=1\cdots n)$ であるとき，共分散行列 Σ を λ_i と $\boldsymbol{\varphi}_i$ を用いて表しなさい．

$$\Sigma = V\Lambda V^T = \begin{bmatrix}\varphi_1 & \cdots & \varphi_n\end{bmatrix}\begin{bmatrix}\lambda_1 & 0 & 0\\ 0 & \ddots & 0\\ 0 & 0 & \lambda_n\end{bmatrix}\begin{bmatrix}\varphi_1^T\\ \vdots\\ \varphi_n^T\end{bmatrix} = \sum_{i=1}^{n}\lambda_i\varphi_i\varphi_i^T$$