2015/02/04　パターン認識特論演習問題

学籍番号＿＿＿＿＿＿＿＿＿＿＿＿＿　＿＿＿＿＿＿＿＿＿＿＿クラスタ　氏名＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

1.　下記の文章の空欄を埋めなさい．（数式 a）~h）1 問 2 点，それ以外 1 問 1 点，合計 26 点）

- パターン$\boldsymbol{x}$がクラスω_j に属すると判定する事によって生じる期待損失$L(\omega_j|\boldsymbol{x})$は，損失関数$l(\omega_i|\omega_j)$を用いて，

$$L(\omega_j|\boldsymbol{x}) = \sum_{i=1}^{C} l(\omega_i|\omega_j)\boxed{\text{a)}\quad\quad}$$

と表現できる．但し，Cはクラス数である．0－1 損失基準では，損失関数$l(\omega_i|\omega_j)$は，

$$l(\omega_i|\omega_j) = \begin{cases} \boxed{\text{b)}} & , i = j \\ & , otherwise \end{cases}$$

と与えられる．このもとでの損失は，

$$L(\omega_j|\boldsymbol{x}) = \sum_{i=1}^{C} l(\omega_i|\omega_j)P(\omega_i|\boldsymbol{x}) = 1 - \boxed{\text{c)}\quad\quad}$$

と表されることから，事後確率を最大化させるクラスに分類すれば損失が最小化できると言える．

- 各クラスの事後確率は，事前生起確率$P(\omega_i)$と，クラス ω_iに属するパターン$\boldsymbol{x}$が生起する確率$P(\boldsymbol{x}|\omega_i)$から$\boxed{\quad\quad}$の定理を用いて計算することができる．しかし，実数ベクトル$\boldsymbol{x}$について$P(\boldsymbol{x}|\omega_i)$を確率の形式で求めることは困難であるため，確率密度関数$p(\boldsymbol{x}|\omega_i)$で代替することが一般的に行われている．

$$\boxed{\text{d)}\quad\quad\quad\quad\quad\quad\quad\quad\quad\quad}$$

- 単純類似度法では，入力パターン$\boldsymbol{x}$と各クラスを代表するパターン$\boldsymbol{r}_i$との余弦を計算し，その値が最も大きいクラスに分類するという計算が行われる．この時の類似度$s(\boldsymbol{x}, \boldsymbol{r}_i)$は，

$$s(\boldsymbol{x}, \boldsymbol{r}_i) = \frac{\boxed{\text{e)}\quad\quad}}{\|\boldsymbol{x}\|\|\boldsymbol{r}_i\|}$$

となるが，クラスに依存しない$\boxed{\text{f)}\quad}$の部分は 1 と置いても良い．

これを拡張した$\boxed{\quad\quad}$類似度法では，クラスiの自己相関行列のj番目の固有ベクトル$\boldsymbol{\varphi}_{ij}(\|\boldsymbol{\varphi}_{ij}\| = 1)$から成る$d$個の代表パターンとの内積計算を行い，それを対応する固有値λ_{ij}によって重み付けした値が類似度として用いられる．パターン$\boldsymbol{x}$のクラスiとの類似度$s_i(\boldsymbol{x})$は，下記のように定義される．

$$s_i(\boldsymbol{x}) = \sum_{j=1}^{d} \frac{\lambda_{ij}\boxed{\text{g)}\quad\quad}}{\lambda_{i1}\|\boldsymbol{x}\|^2}$$

この場合も，$\boxed{\text{h)}\quad}$の部分は 1 と置いても良い．

この計算は$\boxed{\quad\quad\quad\quad}$で行われている射影成分の大小で分類を行う方法と本質的に等価である．

- これ以外に，$\boxed{\quad\quad\quad\quad}$で行われているように，入力に最も近い既知のパターン（プロトタイプ）が帰属するクラスに分類する方法もある．この場合には，各クラスへの帰属度は計算されない．この方法において，$\boxed{\quad\quad\quad\quad\quad}$という手法を用いると，不要なプロトタイプを削除することができる．この際に残されるプロトタイプは，識別境界付近の誤識別を起こしやすい特殊例となる．このほかにも，$\boxed{\quad\quad\quad}$グラフや，RNG を用いたプロトタイプの削減法がある．

- このように，特殊例だけを記憶し，それによって識別を行う手法として$\boxed{\quad\quad}$がある．この方法ではマージン最大化基準に基づいて線形識別面を求めるために，$\boxed{\quad\quad\quad\quad\quad\quad}$を用いた最適化計算を行っているが，この計算によって未定係数αが非ゼロとなるパターンは特殊例すなわち$\boxed{\quad\quad\quad\quad}$となる．

- この例以外にも$\boxed{\quad\quad\quad\quad}$では，誤識別をおこしたトレーニングデータの重みを増して識別器のトレーニングを行いながら，識別器の系列を発生させ，これらの線形結合によって識別器の出力を得ている．このことは，誤識別を起こしやすい部分の重みを増すという意味で，特殊例を重視した識別法であると言える．

2. d次元空間中にn個の独立なベクトルが存在する．これらのベクトルを用いて自己相関行列を求めたとき，そのランクはどれだけになるか？dとnを用いて表わしなさい．（4 点）

3. 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，$\boldsymbol{\varphi}_k^T \Sigma \boldsymbol{\varphi}_k$を求めなさい．（５点）

4. $f(x)$の Fourier 変換を$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx$とする．このとき，次の問いに答えなさい．（4+4+4+3＝15 点）

（１）$f(x)$が平行移動したパターン$g(x) = f(x - \mu)$の Fourier 変換$G(\omega)$を$F(\omega)$を用いた式で表しなさい.

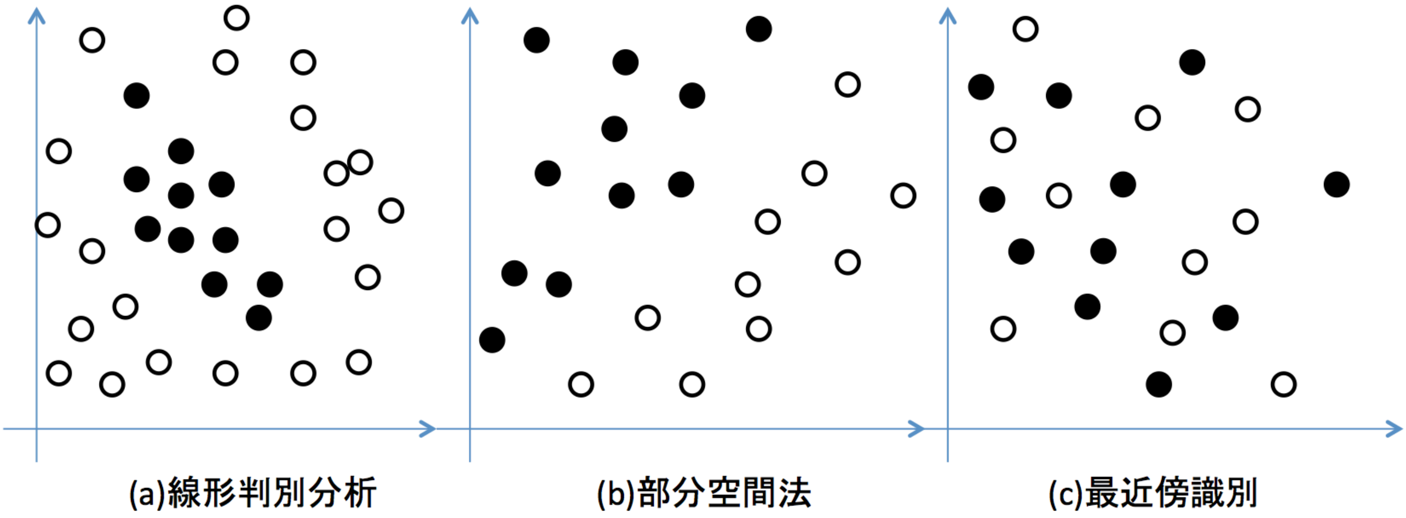
（２）$F(\omega)$とその複素共役を用いてパワースペクトル$\|F(\omega)\|^2$を表しなさい.

（３）$\|G(\omega)\|^2 = \|F(\omega)\|^2$であることを示しなさい.

（４）パワースペクトルによって平行移動に対して不変な特徴を求めることの問題点について説明しなさい.

5. SVM や ADA Boosting が漢字認識に適用しにくい理由を説明しなさい．（５点）

6. 下記の 2 次元特徴分布それぞれに対して用意された識別方法の中で不適切なものを選び，適切な識別器と，その理由を述べよ．（10 点）



7. カーネル主成分分析は，通常の主成分分析で非正則な共分散行列の固有値問題を解かなければならなくなるケースでも，この問題を回避できるという利点がある．その理由について説明しなさい．（5 点）

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

また，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（$i=1,\cdots,n$）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．（5+5=10 点）

9.　直交展開における各基底関数の係数を内積によって決定することが，最小二乗法として妥当である事を説明した下記の文の空白を埋めなさい．（ a)～h)，順に 3+3+3+3+2+2+2+2=20 点）

未知の関数$f(x)$を近似する関数$f^*(x)=\sum_{i=1}^{N}c_i\varphi_i(x)$を考える．このとき，$\int_a^b\left(f(x)-f^*(x)\right)^2dx$の値を最小化する$c_i$を求める問題が最小二乗法である．ここで，以降の準備のために関数$f(x)$と$g(x)$の内積を

$$<f(x),g(x)>=\int_a^b f(x)g(x)dx$$

と定義しておく．この定義に基づき，関数$f(x)$のノルムも以下のように定義できる．

$$\|f(x)\|=\sqrt{<f(x),f(x)>}$$

これらの定義に基づき，最小二乗法において最小化する関数を表現すると，

$$\int_a^b\left(f(x)-f^*(x)\right)^2dx=\|f(x)-f^*(x)\|^2=\left\|f(x)-\sum_{i=1}^{N}c_i\varphi_i(x)\right\|^2$$

$$=\boxed{\text{a)}\qquad\qquad\qquad\qquad}$$

となる．これをc_kで偏微分して 0 とおく事により，目的関数の最小化を行う．

$$\frac{\partial}{\partial c_k}\left(\boxed{\text{a)}\qquad\qquad\qquad\qquad}\right)=\boxed{\text{b)}\qquad\qquad\qquad\qquad}=0$$

すなわち，

$$\sum_{i=1}^{N}\boxed{\text{c)}\qquad\qquad}=\boxed{\text{d)}\qquad\qquad}$$

右の結果を得る．これを連立方程式で表せば，

$$\left[\boxed{\text{e)}\qquad\qquad\qquad\qquad}\right]\begin{bmatrix}c_1\\ \vdots\\ c_N\end{bmatrix}=\left[\boxed{\text{f)}\qquad\qquad}\right]$$

となる．ここで，$\varphi_1(x),\cdots,\varphi_N(x)$が，定義した内積に関して正規直交基底をなす場合，

$$\begin{bmatrix}c_1\\ \vdots\\ c_N\end{bmatrix}=\left[\boxed{\text{g)}\qquad\qquad}\right]$$

となることから，

$$f^*(x)=\sum_{i=1}^{N}\boxed{\text{h)}\qquad\qquad}\varphi_i(x)$$

と表すことが出来，最小二乗法の結果と直交展開の結果は一致することが判る．

2015/02/04　パターン認識特論演習問題解答例

学籍番号＿＿＿＿＿＿＿＿ ＿＿＿＿＿＿＿クラスタ　氏名＿＿＿＿＿＿＿＿＿＿＿＿＿＿＿

1.　下記の文章の空欄を埋めなさい．（数式 1 問 2 点，それ以外 1 問 1 点，合計 26 点）

- パターン$\boldsymbol{x}$がクラスω_j に属すると判定する事によって生じる期待損失$L(\omega_j|\boldsymbol{x})$は，損失関数$l(\omega_i|\omega_j)$を用いて，

$$L(\omega_j|\boldsymbol{x}) = \sum_{i=1}^{C} l(\omega_i|\omega_j)\boxed{P(\omega_i|\boldsymbol{x})}$$

と表現できる．但し，Cはクラス数である．0−1 損失基準では，損失関数$l(\omega_i|\omega_j)$は，

$$l(\omega_i|\omega_j) = \begin{cases} \boxed{0} & ,i = j \\ \boxed{1} & ,otherwise \end{cases}$$

と与えられる．このもとでの損失は，

$$L(\omega_j|\boldsymbol{x}) = \sum_{i=1}^{C} l(\omega_i|\omega_j)P(\omega_i|\boldsymbol{x}) = 1 - \boxed{P(\omega_j|\boldsymbol{x})}$$

と表されることから，事後確率を最大化させるクラスに分類すれば損失が最小化できると言える．

- 各クラスの事後確率は，事前生起確率$P(\omega_i)$と，クラス ω_iに属するパターン$\boldsymbol{x}$が生起する確率$P(\boldsymbol{x}|\omega_i)$から Bayes の定理を用いて計算することができる．しかし，実数ベクトル$\boldsymbol{x}$について$P(\boldsymbol{x}|\omega_i)$を確率の形式で求めることは困難であるため，確率密度関数$p(\boldsymbol{x}|\omega_i)$で代替することが一般的に行われている．

$$\boxed{P(\omega_i|\boldsymbol{x}) = \frac{P(\omega_i)p(x|\omega_i)}{\sum_{j=1}^{C} P(\omega_j)p(x|\omega_j)}}$$

- 単純類似度法では，入力パターン$\boldsymbol{x}$と各クラスを代表するパターン$\boldsymbol{r}_i$との余弦を計算し，その値が最も大きいクラスに分類するという計算が行われる．この時の類似度$s(\boldsymbol{x},\boldsymbol{r}_i)$は，

$$s(\boldsymbol{x},\boldsymbol{r}_i) = \frac{\boxed{\boldsymbol{x}^T\boldsymbol{r}_i}}{\|\boldsymbol{x}\|\|\boldsymbol{r}_i\|}$$

となるが，クラスに依存しない$\boxed{\|\boldsymbol{x}\|}$の部分は 1 と置いても良い．
これを拡張した 複合 類似度法では，クラス$\boldsymbol{i}$の自己相関行列のj番目の固有ベクトル$\boldsymbol{\varphi}_{ij}(\|\boldsymbol{\varphi}_{ij}\| = 1)$から成る$d$個の代表パターンとの内積計算を行い，それを対応する固有値λ_{ij}によって重み付けした値が類似度として用いられる．パターン$\boldsymbol{x}$のクラス$\boldsymbol{i}$との類似度$s_i(\boldsymbol{x})$は，下記のように定義される．

$$s_i(\boldsymbol{x}) = \sum_{j=1}^{d} \frac{\lambda_{ij}\boxed{(\boldsymbol{x}^T\boldsymbol{\varphi}_{ij})^2}}{\lambda_{i1}\|\boldsymbol{x}\|^2}$$

この場合も，$\boxed{\|\boldsymbol{x}\|}$の部分は 1 と置いても良い．
この計算は 部分空間法 で行われている射影成分の大小で分類を行う方法と本質的に等価である．

- これ以外に， 最近傍識別器 で行われているように，入力に最も近い既知のパターン（プロトタイプ）が帰属するクラスに分類する方法もある．この場合には，各クラスへの帰属度は計算されない．この方法において， Voronoi Condensing という手法を用いると，不要なプロトタイプを削除することができる．この際に残されるプロトタイプは，識別境界付近の誤識別を起こしやすい特殊例となる．このほかにも， Gabriel グラフや，RNG を用いたプロトタイプの削減法がある．
- このように，特殊例だけを記憶し，それによって識別を行う手法として SVM がある．この方法ではマージン最大化基準に基づいて線形識別面を求めるために， ラグランジェの未定係数法 を用いた最適化計算を行っているが，この計算によって未定係数αが非ゼロとなるパターンは特殊例すなわち サポートベクター となる．
- この例以外にも ADA Boosting では，誤識別をおこしたトレーニングデータの重みを増して識別器のトレーニングを行いながら，識別器の系列を発生させ，これらの線形結合によって識別器の出力を得ている．このことは，誤識別を起こしやすい部分の重みを増すという意味で，特殊例を重視した識別法であると言える．

2. d次元空間中にn個の独立なベクトルが存在する．これらのベクトルを用いて自己相関行列を求めたとき，そのランクはどれだけになるか？dとnを用いて表わしなさい．（4 点）

$$\min(d, n)$$

3. 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，$\boldsymbol{\varphi}_k^T \Sigma \boldsymbol{\varphi}_k$を求めなさい．（５点）

$$\Sigma = \sum_{i=1}^{n} \lambda_i \boldsymbol{\varphi}_i \boldsymbol{\varphi}_i^T = [\boldsymbol{\varphi}_1 \quad \cdots \quad \boldsymbol{\varphi}_n] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\varphi}_1^T \\ \vdots \\ \boldsymbol{\varphi}_n^T \end{bmatrix}$$

したがって，

$$\boldsymbol{\varphi}_k^T \Sigma \boldsymbol{\varphi}_k = \boldsymbol{\varphi}_k^T [\boldsymbol{\varphi}_1 \quad \cdots \quad \boldsymbol{\varphi}_n] \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{bmatrix} \begin{bmatrix} \boldsymbol{\varphi}_1^T \\ \vdots \\ \boldsymbol{\varphi}_n^T \end{bmatrix} \boldsymbol{\varphi}_k = \lambda_k$$

4. $f(x)$の Fourier 変換を$F(\omega) = \int_{-\infty}^{\infty} f(x) e^{-j\omega x} dx$とする．このとき，次の問いに答えなさい．（4+4+4+3＝15 点）

（１）$f(x)$が平行移動したパターン$g(x) = f(x - \mu)$の Fourier 変換$G(\omega)$を$F(\omega)$を用いた式で表しなさい．

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

（２）$F(\omega)$とその複素共役を用いてパワースペクトル$\|F(\omega)\|^2$を表しなさい．

$$\|F(\omega)\|^2 = F^*(\omega) F(\omega)$$

但し，$F^*(\omega)$は$F(\omega)$の複素共役である．

（３）$\|G(\omega)\|^2 = \|F(\omega)\|^2$であることを示しなさい．

$$\begin{aligned} \|G(\omega)\|^2 &= G^*(\omega) G(\omega) = \left(e^{-j\omega\mu} F(\omega)\right)^* e^{-j\omega\mu} F(\omega) = e^{j\omega\mu} F^*(\omega) e^{-j\omega\mu} F(\omega) = e^{j(\omega\mu - \omega\mu)} F^*(\omega) F(\omega) \\ &= \|F(\omega)\|^2 \end{aligned}$$

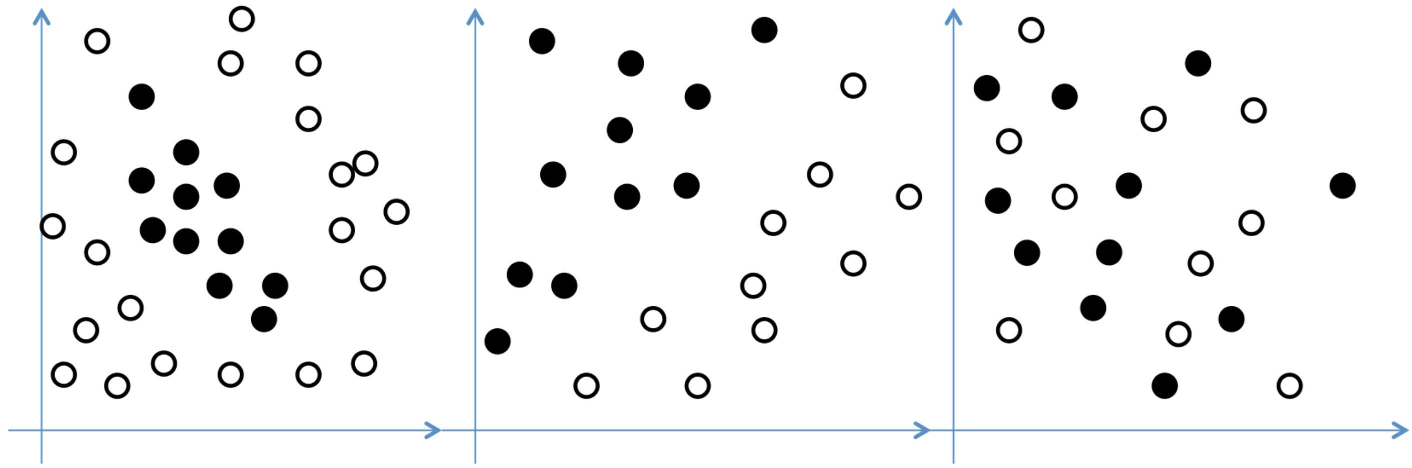
（４）パワースペクトルによって平行移動に対して不変な特徴を求めることの問題点について説明しなさい．

パワースペクトルを用いて平行移動成分を除去すると，各ωにおける位相情報（実部と虚部の比の逆正接）も失われてしまう．離散 Fourier 変換を行う場合は，独立な情報は半分になってしまう．平行移動は単一のパラメータμに起因する変化であるにもかかわらず，特徴ベクトルの半分を失うことは，識別に有効な特徴まで失われていることを意味している．

5. SVM や ADA Boosting が漢字認識に適用しにくい理由を説明しなさい．（５点）

SVM や ADA Boosting は基本的に２クラスの識別問題に適用される判別モデルである．したがって，クラス数Nが極めて多い問題では，各クラス間の隣接関係を調べる必要がある．このチェックのためには ${}_N C_2$組の組み合わせを調べなければならない．さらに，識別時には決定木を構築し，どの順序で識別を行うべきかを考えなければならない．したがって，Nが数万のオーダーになる漢字の文字認識などには適さないと言える．

6.　下記の 2 次元特徴分布それぞれに対して用意された識別方法の中で不適切なものを選び，適切な識別器と，その理由を述べよ．（10 点）



(a)線形判別分析　　(b)部分空間法　　(c)最近傍識別

(a)の線形判別分析が誤り．理由：線形判別分析は 1 次元の軸にパターンを射影し，その軸上のクラス内分散に対するクラス間分散の比を最大化する手法である．したがって，2 次元平面上で一方が他方を取り囲むような分布は，1 次元の軸に射影しても分離することは出来ない．このような場合は，取り囲んでいる○の方のパターンに関しては，適切な分布が求まらないケースが多いため，判別モデルに基づく手法が適している．SVM, ADA Boosting,最近傍識別などが適用可能である．

7.　カーネル主成分分析は，通常の主成分分析で非正則な共分散行列の固有値問題を解かなければならなくなるケースでも，この問題を回避できるという利点がある．その理由について説明しなさい．（5 点）

カーネル主成分分析では，主軸はトレーニングデータの線形結合で表現されることを仮定し，トレーニングデータ間のカーネル関数を要素とする「グラム行列」の固有値問題を解くことで，この線形結合係数を求めるという計算が行われる．このため，次元数＋1 よりも少ないサンプルしか与えられない場合でも，グラム行列が正則であれば，安定に固有値問題を解くことができる．

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

また，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（$i = 1, \cdots, n$）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．（5+5=10 点）

$$P = UU^T$$

$$\|\boldsymbol{x} - P_i\boldsymbol{x}\|^2 = (\boldsymbol{x} - P_i\boldsymbol{x})^T(\boldsymbol{x} - P_i\boldsymbol{x}) = \|\boldsymbol{x}\|^2 - 2\boldsymbol{x}^T P_i \boldsymbol{x} + \boldsymbol{x}^T {P_i}^T P_i \boldsymbol{x}$$

となる．$P_i = U_i {U_i}^T$ から $P_i^T P_i = P_i$ となるので，結局$\|\boldsymbol{x} - P_i\boldsymbol{x}\|^2 = \|\boldsymbol{x}\|^2 - \boldsymbol{x}^T P_i \boldsymbol{x}$となる．このうち，$\boldsymbol{x}$は$i$に依存せず最小化とは無関係であるので，$\|\boldsymbol{x}\|^2$を無視することができ，$-\boldsymbol{x}^T P_i \boldsymbol{x}$を最小化すればよいことが分かる．したがって，$\|\boldsymbol{x} - P_i\boldsymbol{x}\|^2$を最小化する問題は，$\boldsymbol{x}^T P_i \boldsymbol{x}$を最大化する問題となっている．

9.　直交展開における各基底関数の係数を内積によって決定することが，最小二乗法として妥当である事を説明した下記の文の空白を埋めなさい．（ a)～h)，順に 3+3+3+3+2+2+2+2=20 点）

未知の関数$f(x)$を近似する関数$f^*(x)=\sum_{i=1}^{N} c_i\varphi_i(x)$を考える．このとき，$\int_a^b \left(f(x)-f^*(x)\right)^2 dx$の値を最小化する$c_i$を求める問題が最小二乗法である．ここで，以降の準備のために関数$f(x)$と$g(x)$の内積を

$$< f(x), g(x) >= \int_a^b f(x)g(x)dx$$

と定義しておく．この定義に基づき，関数$f(x)$のノルムも以下のように定義できる．

$$\|f(x)\| = \sqrt{< f(x), f(x) >}$$

これらの定義に基づき，最小二乗法において最小化する関数を表現すると，

$$\int_a^b \left(f(x)-f^*(x)\right)^2 dx = \|f(x)-f^*(x)\|^2 = \left\| f(x) - \sum_{i=1}^{N} c_i\varphi_i(x) \right\|^2$$

$$= \boxed{\text{a)}\ \|f(x)\|^2 - 2 < f(x), \sum_{i=1}^{N} c_i\varphi_i(x) > + \left\| \sum_{i=1}^{N} c_i\varphi_i(x) \right\|^2}$$

となる．これをc_kで偏微分して 0 とおく事により，目的関数の最小化を行う．

$$\frac{\partial}{\partial c_k}\left(\boxed{\text{a)}\ \|f(x)\|^2 - 2 < f(x), \sum_{i=1}^{N} c_i\varphi_i(x) > + \left\| \sum_{i=1}^{N} c_i\varphi_i(x) \right\|^2} \right) = \boxed{\text{b)}\ -2 < f(x), \varphi_k(x) > +2 < \varphi_k(x), \sum_{i=1}^{N} c_i\varphi_i(x) >} = 0$$

すなわち，

$$\sum_{i=1}^{N} \boxed{\text{c)}\ c_i < \varphi_k(x), \varphi_i(x) >} = \boxed{\text{d)}\ < f(x), \varphi_k(x) >}$$

右の結果を得る．これを連立方程式で表せば，

$$\left[\boxed{\text{e)}\ \begin{matrix} < \varphi_1(x), \varphi_1(x) > & \dots & < \varphi_1(x), \varphi_N(x) > \\ \vdots & \ddots & \vdots \\ < \varphi_N(x), \varphi_1(x) > & \dots & < \varphi_N(x), \varphi_N(x) > \end{matrix}}\right] \begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} = \left[\boxed{\text{f)}\ \begin{matrix} < \varphi_1(x), f(x) > \\ \vdots \\ < \varphi_N(x), f(x) > \end{matrix}}\right]$$

となる．ここで，$\varphi_1(x), \cdots, \varphi_N(x)$が，定義した内積に関して正規直交基底をなす場合，

$$\begin{bmatrix} c_1 \\ \vdots \\ c_N \end{bmatrix} = \left[\boxed{\text{g)}\ \begin{matrix} < \varphi_1(x), f(x) > \\ \vdots \\ < \varphi_N(x), f(x) > \end{matrix}}\right]$$

となることから，

$$f^*(x) = \sum_{i=1}^{N} \boxed{\text{h)}\ < \varphi_i(x), f(x) >} \varphi_i(x)$$

と表すことが出来，最小二乗法の結果と直交展開の結果は一致することが判る．