

カリキュラム ('05) ('15)	データ解析入門	担当教員：内山俊郎	2単位
設 題			
課題 1	<p>クラスタリングとベクトル量子化の定義を、その違いが分かるように説明し、平方和最小基準クラスタリング問題と 2 乗誤差に基づくベクトル量子化問題の局所最適解における性質について記述しなさい。</p>		
課題 2	<p>教科書の「HokkaidoTowns_xy.dat」データを k-means アルゴリズムと競合学習で、5 分割にクラスタリングする実験を行い、その結果を示し、考察しなさい。 具体的には、プログラムに与える引数である「乱数の種」、「学習率」（競合学習のみ）などをいくつか変更して実験を行い、両手法の性能比較を行いなさい。</p>		
課題 3	<p>教科書の「HokkaidoTowns_xy.dat」データを、k-means アルゴリズムなどで 5 分割したクラスタリングを行い、その結果を「Hokkaido_xyl.dat」としなさい。 次に、修正パーセプトロン学習プログラムである modPerceptron を用い、上記クラスタリング結果を学習パターンとして学習させなさい。その時、「乱数の種」、「学習率」を変化させて実験し、収束するまでの繰り返し回数の傾向について考察しなさい。</p>		
課題 4	<p>教科書の「HokkaidoTowns_xy.dat」の各カラムが x y 座標の座標値を表すとします。このデータの重みベクトルによるボロノイ分割が、y 軸を境界とする 2 分割になりました。この重みベクトルはどのようなものになるか例を使って説明しなさい。</p>		
課題 5	<p>赤と青の 2 つの箱があり、赤の箱にはりんごが 2 個とオレンジが 6 個、青の箱にはりんごが 3 個とオレンジが 1 個入っているとします。箱の 1 つをランダムに選び、果物をランダムに 1 個取り出す。その際、赤の箱を 40%、青の箱を 60% 選び、箱の中の果物は別け隔てなく同じ確からしさで選ぶ。この時、オレンジを選び出したとして、それが青い箱からとり出されたものである確率はいくらか？ 導出過程も含めて書きなさい。</p>		
課題 6	<p>教科書 5 章の表 5.2 「フルーツのブレンド割合」の値は、壺を選んだ時に各フルーツが取り出される条件付き確率を表しています。今、各壺が選ばれる確率がすべて等しいとします。この時、下記のフルーツポンチがあるとき、それぞれの壺が選ばれたという「対数事後確率の大小関係を比較できる値」を求め比較しなさい。 フルーツポンチ：マンゴー 1 個、ナタデココ 2 個、ぶどう 3 個、さくらんぼ 5 個</p>		
課題 7	<p>フルーツポンチを genPunch で生成し、それを estParam で処理して、モデルのパラメータを推定する実験をしなさい。 ボウルの数や果物の数を変えると、推定結果の精度にどのような影響があるかについて考察しなさい。</p>		
作成方法は「ワープロ（推奨）」又は「筆記」			
ワープロ	用紙等：本学標準フォーマット・コピー用紙等（無地）		
筆 記	筆記用具：特に指定しない		
	用紙：市販の A4 レポート用紙		
文字数等	文字数の制限はしない・横書き		