Министерство общего и профессионального образования РФ

Российская Академия Наук

Самарский государственный аэрокосмический университет
имени академика С.П. Королева

Институт систем обработки изображений РАН

МОДЕЛИРОВАНИЕ ЭКСПЕРИМЕНТАЛЬНЫХ ДАННЫХ ДЛЯ РЕШЕНИЯ ЗАДАЧ РАСПОЗНАВАНИЯ ОБРАЗОВ

Методические указания к лабораторной работе № 1
по курсу «МЕТОДЫ РАСПОЗНАВАНИЯ ОБРАЗОВ»

САМАРА
2000

Составители: к.ф.-м.н., доцент Э.И.Коломиец

к.т.н. В.В.Мясников

УДК 681.3

Моделирование экспериментальных данных
для решения задач распознавания образов

Методические указания к лабораторной работе № 1
Самарский государственный аэрокосмический университет
имени академика С.П.КОРОЛЕВА
Составители: Э.И. Коломиец, В.В. Мясников
Самара, 2000. 21 с.

В лабораторной работе № 1 по курсу «Методы распознавания образов» изучаются методы получения выборочных данных, являющихся реализациями нормально распределенных случайных векторов и бинарных случайных векторов с независимыми координатами. Эти данные предназначены для изучения в последующем различных методов и алгоритмов классификации. Дается описание среды математического программирования MATLAB, в рамках которой выполняется лабораторная работа.

Методические указания предназначены для студентов специальности 01.02 “Прикладная математика”, обучающихся по специализации 01.02.01 «Математическое и программное обеспечение обработки изображений».

Печатается по решению редакционно-издательского совета
Самарского государственного аэрокосмического университета
имени академика С.П. Королева

Рецензент: д.ф.-м.н., профессор А.И.Жданов

Данные методические указания подготовлены при поддержке
Федеральной целевой программы “Государственная поддержка
интеграции высшего образования и фундаментальной науки на 1997-2000 гг.”
(Постановление Правительства РФ №1062 от 09.09.96)

Цель работы - подготовка экспериментального материала для решения задач распознавания образов, получение навыков работы в среде MATLAB.

1. ТЕОРЕТИЧЕСКИЕ ОСНОВЫ ЛАБОРАТОРНОЙ РАБОТЫ

1.1. Моделирование случайного вектора с нормальным законом распределения

Пусть $\overline{X} = (X_0, \ldots, X_{n-1})^T$ - $n-$мерный случайный вектор, имеющий нормальный закон распределения: $\overline{X} \sim N(\overline{M}, B)$. Это означает, что плотность вероятностей случайного вектора $\overline{X}$ имеет вид:

$$f(\overline{x}) = \frac{1}{(2\pi)^{n/2}\sqrt{|B|}} \exp\left(-\frac{1}{2}(\overline{x} - \overline{M})^T B^{-1} (\overline{x} - \overline{M})\right),$$

где $|\ldots|$ - определитель матрицы, $(\ldots)^T$ - транспонирование матрицы (вектора). Вектор $\overline{M} = (M_0, \ldots, M_{n-1})^T$ представляет собой вектор математических ожиданий координат вектора $\overline{X}$: $M_i = MX_i$ $\left(i = \overline{0, n-1}\right)$; B – корреляционная матрица

$$B = \begin{pmatrix} B_{00} & B_{01} & \ldots & B_{0(n-1)} \\ B_{10} & B_{11} & \ldots & B_{1(n-1)} \\ \vdots & & & \\ B_{(n-1)0} & B_{(n-1)1} & \ldots & B_{(n-1)(n-1)} \end{pmatrix},$$

элементами которой являются всевозможные корреляционные моменты: $B_{ij} = M(X_i - M_i)(X_j - M_j)$, $\left(i, j = \overline{0, n-1}\right)$. Очевидно, что вектор $\overline{M}$ и матрица B полностью определяют нормальный закон распределения.

Вектор $\overline{X} \sim N(\overline{M}, B)$ можно получить специальным линейным преобразованием вектора $\overline{\xi} = (\xi_0, \ldots, \xi_{n-1})^T$, компоненты которого суть независимые случайные величины, имеющие стандартный нормальный закон распределения:

$$\xi_i \sim N(0, 1), \quad m_i = 0, \quad \sigma_i^2 = 1, \quad \left(i = \overline{0, n-1}\right).$$

Обычно предполагают, что матрица A преобразования $\overline{X} = A\overline{\xi} + \overline{M}$ является треугольной, то есть

$$A = \begin{pmatrix} a_{00} & 0 & \dots & 0 \\ a_{10} & a_{11} & \dots & 0 \\ \vdots & & & \\ a_{(n-1)0} & a_{(n-1)1} & \dots & a_{(n-1)(n-1)} \end{pmatrix}.$$

Коэффициенты a_{ij} легко определяются рекуррентным образом. Действительно, для диагональных элементов матрицы B справедливо соотношение:

$$B_{ii} = M[X_i - M_i]^2 = M\left[\left(\sum_{k=0}^{i} a_{ik}\,\xi_k + M_i\right) - M_i\right]^2 = M\sum_{k=0}^{i}\ \sum_{l=0}^{i} a_{ik}\,a_{il}\,\underbrace{\xi_k \xi_l}_{\delta_{kl}} = \sum_{k=0}^{i} a_{ik}^2 = a_{ii}^2 + \sum_{k=0}^{i-1} a_{ik}^2 \quad ,$$

откуда

$$a_{ii} = \sqrt{B_{ii} - \sum_{k=0}^{i-1} a_{ik}^2}, \quad i = \overline{1, n-1}; \qquad a_{00} = \sqrt{B_{00}}\ .$$

Для недиагональных элементов матрицы B выполняется равенство:

$$B_{ij} = M(X_i - M_i)(X_j - M_j) = M\left(\sum_{k=0}^{i} a_{ik}\,\xi_k\right)\left(\sum_{l=0}^{j} a_{jl}\,\xi_l\right) = \sum_{k=0}^{i}\ \sum_{l=0}^{j} a_{ik} a_{jl}\, \underbrace{M\xi_k \xi_l}_{\delta_{kl}}\ .$$

Предполагая, что $i < j$, получаем:

$$B_{ij} == \sum_{k=0}^{i} a_{ik}\,a_{jk} \ = \sum_{k=0}^{i-1} a_{ik} a_{jk} + a_{ii} a_{ji}\ ,$$

откуда

$$a_{ji} = \frac{1}{a_{ii}}\left(B_{ij} - \sum_{k=0}^{i-1} a_{ik} a_{jk}\right), \quad 1 \le i < j \le n-1; \qquad a_{j0} = \frac{B_{0j}}{a_{00}}, \quad j = \overline{1, n-1};.$$

В частном случае, когда случайный вектор является двумерным (n=2), получаем следующие выражения для элементов матрицы преобразования:

$$a_{00} = \sqrt{B_{00}}, \quad a_{21} = \frac{B_{01}}{\sqrt{B_{00}}}, \quad a_{11} = \sqrt{B_{11} - \frac{B_{01}^2}{B_{00}}}$$

Заметим, что поскольку для элементов матрицы B справедливо неравенство $\left|B_{ij}\right| \le \sqrt{B_{ii}\,B_{jj}}\ ,\quad \left(i, j = \overline{0, n-1}\right)$, то все коэффициенты a_{ij} корректно определены в том смысле, что подкоренные выражения в приведенных соотношениях всегда неотрицательны.

1.2. Оценивание параметров нормального закона распределения

Если n-мерный случайный вектор $\overline{X}$ имеет нормальный закон распределения $N(\overline{M}, B)$, то оценки максимального правдоподобия его математического ожидания $\hat{\overline{M}}$ и корреляционной матрицы $\hat{B}$, найденные по выборке $\overline{x}_1,\ldots,\overline{x}_N$ объема N, выглядят следующим образом:

$$\hat{\overline{M}} = \frac{1}{N}\sum_{i=1}^{N}\overline{x}_i \ , \qquad \hat{B} = \frac{1}{N}\sum_{i=1}^{N}\left(\overline{x}_i - \hat{\overline{M}}\right)\left(\overline{x}_i - \hat{\overline{M}}\right)^T = \frac{1}{N}\sum_{i=1}^{N}\overline{x}_i\overline{x}_i^T - \hat{\overline{M}}\ \hat{\overline{M}}^T \ .$$

Оценки вектора средних $\hat{\overline{M}}$ и корреляционной матрицы $\hat{B}$ можно задать и рекуррентными соотношениями, когда коррекция их значений, вычисленных по выборке объема N, производится с учетом появления каждого нового элемента выборки. Обозначим $\hat{\overline{M}}(N)$ и $\hat{B}(N)$ - оценки $\hat{\overline{M}}$ и $\hat{B}$, вычисленные по выборке объема N. Полагая на первом шаге $\hat{\overline{M}}(1) = \overline{x}_1$, имеем

$$\hat{\overline{M}}(N+1) = \frac{1}{N+1}\sum_{i=1}^{N+1}\overline{x}_i = \frac{1}{N+1}\left(N\,\hat{\overline{M}}(N) + \overline{x}_N\right) \ .$$

Пополнение выборки одним элементом при расчете оценки корреляционной матрицы приводит к следующему выражению:

$$\hat{B}(N+1) = \frac{1}{N+1}\sum_{i=1}^{N+1}\overline{x}_i\overline{x}_i^T - \hat{\overline{M}}(N+1)\hat{\overline{M}}^T(N+1) =$$

$$= \frac{1}{N+1}\left(N\ \hat{B}(N) + N\,\hat{\overline{M}}(N)\hat{\overline{M}}^T(N+1) + \overline{x}_{N+1}\overline{x}_{N+1}^T\right) - \frac{1}{(N+1)^2}\left(N\,\hat{\overline{M}}(N) + \overline{x}_{N+1}\right)\left(N\,\hat{\overline{M}}(N) + \overline{x}_{N+1}\right)$$

При этом на первом шаге $\hat{B}(1) = 0$, так как $\hat{\overline{M}}(1) = \overline{x}_1$.

1.3. Меры близости нормальных распределений

Пусть $f_0(\overline{x})$ и $f_1(\overline{x})$ - плотности вероятностей нормально распределенного случайного вектора с параметрами:

$$f_0 \sim N\left(\overline{M}_0, B_0\right) \text{ и } f_1 \sim N\left(\overline{M}_1, B_1\right).$$

Мерой близости распределений $f_0(\overline{x})$ и $f_1(\overline{x})$ является *расстояние Бхатачария*, вычисляемое по формуле:

$$\rho_B = \frac{1}{4}\left(\overline{M}_1 - \overline{M}_0\right)^T \left(\frac{B_1 + B_0}{2}\right)^{-1}\left(\overline{M}_1 - \overline{M}_0\right) + \frac{1}{2}\ln\frac{\left|\dfrac{B_1 + B_0}{2}\right|}{\sqrt{|B_1| \cdot |B_0|}}. \qquad (1)$$

Для случая равных корреляционных матриц $(B_1 = B_0 = B)$ в качестве меры близости распределений используют расстояние Махаланобиса между векторами средних двух нормальных распределений:

$$\rho_M\left(\overline{M}_0, \overline{M}_1\right) = \left(\overline{M}_1 - \overline{M}_0\right)^T B^{-1}\left(\overline{M}_1 - \overline{M}_0\right), \qquad (2)$$

которое в этой ситуации с точностью до постоянного множителя совпадает с расстоянием Бхатачария. Если компоненты случайного вектора $\overline{X}$ независимы и одинаково распределены, то есть корреляционная матрица удовлетворяет условию $B = D_X I$, где I – единичная $N \times N$ матрица, а D_X - дисперсия компонент случайного вектора, то близость нормальных распределений в смысле расстояния Махаланобиса и, соответственно, Бхатачария эквивалентна близости в смысле евклидова расстояния между векторами средних:

$$\rho_E\left(\overline{M}_0, \overline{M}_1\right) = \left\|\overline{M}_1 - \overline{M}_0\right\|^2 = \left(\overline{M}_1 - \overline{M}_0\right)^T\left(\overline{M}_1 - \overline{M}_0\right).$$

Использование метрик Бхатачария или Махаланобиса в общем случае предпочтительнее евклидовой, поскольку они учитывают как дисперсии отдельных компонент случайного вектора, так и их взаимные корреляции.

Расстояния Бхатачария и Махаланобиса обладают следующим важным для задач распознавания свойством.

Утверждение. *Расстояния Бхатачария и Махаланобиса инвариантны относительно любого невырожденного линейного преобразования случайного вектора.*

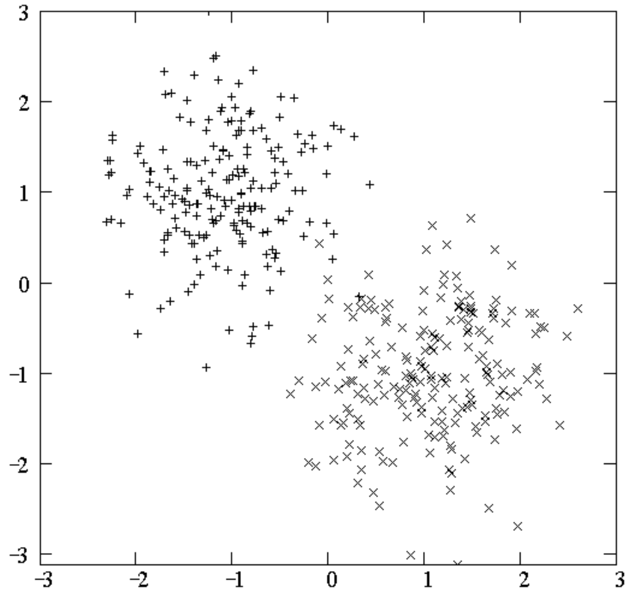
Действительно, пусть вектор $\overline{Y}$ получен в результате линейного преобразования нормально распределенного случайного вектора $\overline{X}$: $\overline{Y} = C\overline{X} + \overline{E}$, где C - матрица преобразования с отличным от нуля определителем $(|C| \neq 0)$, отвечающая за поворот и масштабирование координатных осей, а $\overline{E}$ - вектор, определяющий смещения начала координат. Случайный вектор $\overline{Y}$ оказывается также распределенным нормально с параметрами:

$$\overline{M}_l^Y = C\overline{M}_l + E, \qquad B_l^Y = CB_lC^T, \quad (l = 0{,}1). \tag{3}$$

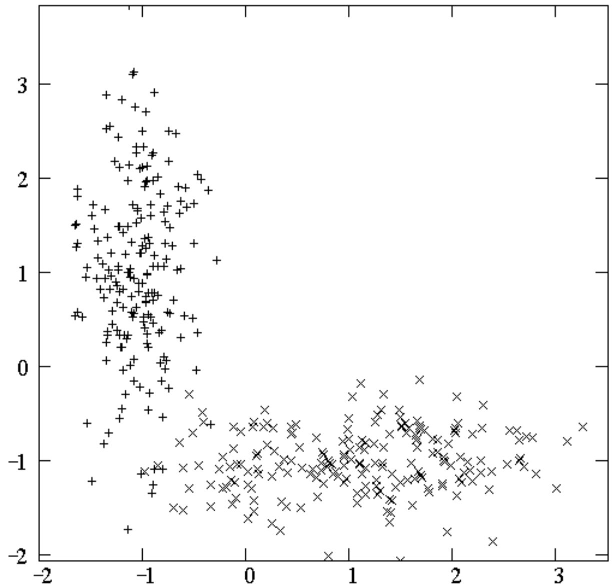
Подставляя (3) в выражения для расстояний (1) и (2) и учитывая справедливость следующих тождеств для произвольных невырожденных матриц C и B:

$$|BC| = |B||C|, \quad (BC)^T = C^T B^T, \quad (BC)^{-1} = C^{-1}B^{-1},$$

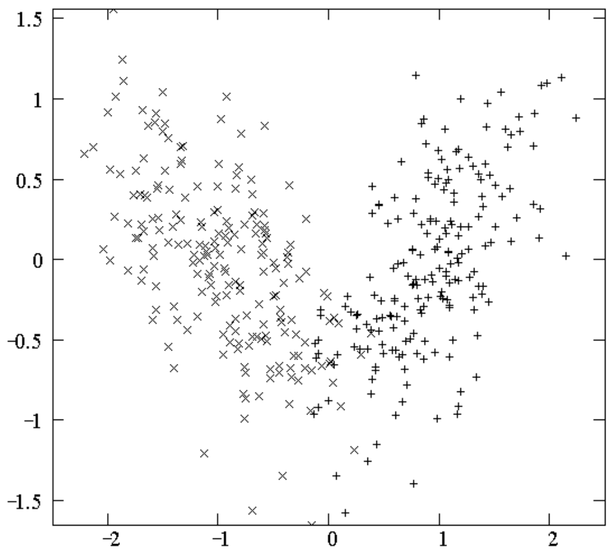
убеждаемся в справедливости приведенного утверждения.



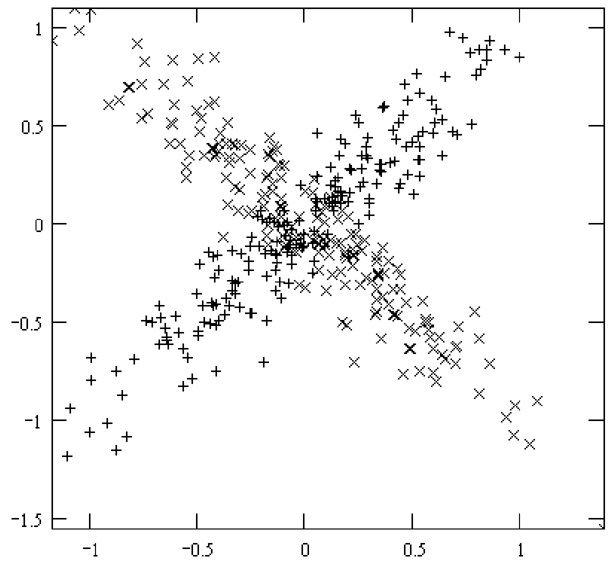
$B_0 = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, $B_1 = \begin{bmatrix} a & 0 \\ 0 & a \end{bmatrix}$, $\rho_B = 5$, $\rho_M = 20$;



$B_0 = \begin{bmatrix} a & 0 \\ 0 & b \end{bmatrix}$, $B_1 = \begin{bmatrix} b & 0 \\ 0 & a \end{bmatrix}$, $\rho_B = 5$, $\rho_M = 18$;



$B_0 = \begin{bmatrix} a & -b \\ -b & a \end{bmatrix}$, $B_1 = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$, $\rho_B = 3.5$, $\rho_M = 14$;



$B_0 = \begin{bmatrix} a & -b \\ -b & a \end{bmatrix}$, $B_1 = \begin{bmatrix} a & b \\ b & a \end{bmatrix}$, $\rho_B = 1$, $\rho_M = 0$.

Рис.1. *Примеры реализаций нормально распределенных случайных векторов*
(f_0 – "×", f_1 – "+", a>b>0)

1.4. Моделирование бинарных случайных векторов с независимыми координатами

Пусть $\overline{X} = (X_0, \ldots, X_{n-1})^T$ - n – мерный бинарный случайный вектор, компоненты которого принимают одно из двух значений $\{0,1\}$. Закон распределения бинарного случайного вектора задается совокупностью вероятностей $P(\overline{X} = \overline{x})$ для всех возможных значений $\overline{x} = (x_0, \ldots, x_{n-1})^T$ вектора. Если координаты вектора $\overline{X}$ независимы, то распределение вероятностей записывается в виде:

$$P(\overline{X} = \overline{x}) = \prod_{i=0}^{n-1} P(X_i = x_i) = \prod_{i=0}^{n-1} \left(p_i x_i + (1 - p_i)(1 - x_i)\right),$$

где $p_i = P(X_i = 1)$. Таким образом, для формирования одной реализации бинарного случайного вектора с независимыми координатами необходимо получить по одной реализации каждой из n бинарных случайных величин X_i $\left(i = 0, n-1\right)$.

Стандартный метод моделирования бинарной случайной величины X с распределением вероятностей $P(X = 1) = p$, $P(X = 0) = 1 - p$ основан на следующих очевидных соотношениях:

$$P\{0 \le U < 1 - p\} = 1 - p, \quad P\{1 - p < U \le 1\} = p,$$

где U – равномерно распределенная на отрезке $[0,1]$ случайная величина: $U \sim R[0,1]$. Таким образом случайные величины U и X связаны соотношением:

$$X = [U/(1 - p)],$$

где $[\ldots]$ - целая часть числа. Следовательно, компоненты одной реализации искомого вектора могут быть получены по формуле:

$$x_i = \lfloor u_i/(1 - p_i) \rfloor, \quad i = \overline{0, n-1},$$

здесь u_i - независимые реализации случайной величины U.

2. СПРАВОЧНЫЕ СВЕДЕНИЯ О СИСТЕМЕ МАТЕМАТИЧЕСКОГО ПРОГРАММИРОВАНИЯ MATLAB

MATLAB – пакет прикладных программ для решения задач технических вычислений и одноимённый язык программирования, используемый в этом пакете. MATLAB используют более 1 000 000 инженерных и научных работников, он работает на большинстве современных операционных систем, включая Linux, Mac OS и Microsoft Windows. Данный программный продукт обеспечивает пользователя всеми необходимыми средствами для быстрого и эффективного решения математических задач и позволяет производить как традиционные численные, так и более сложные аналитические вычисления.

Язык MATLAB является высокоуровневым интерпретируемым языком программирования, включающим основанные на матрицах структуры данных, широкий спектр функций, интегрированную среду разработки, объектно-ориентированные возможности и интерфейсы к программам, написанным на других языках программирования. Основной особенностью языка MATLAB являются его широкие возможности по работе с матрицами, которые создатели языка выразили в лозунге «думай векторно».

Некоторые примеры работы в среде MATLAB приведены на Рис.2.

2.1. Рабочая среда MATLAB

Внешний вид MATLAB[1] как приложения Windows интуитивно понятен. Область приложения разделена на несколько частей (см. Рис.3), включающие в себя меню, панель инструментов, панель ввода команд («Command Window»), панель рабочего пространства («Workspace»), панель истории команд («Command History»), панель рабочего каталога («Current Folder»), где и располагается рабочий файл программы и формируемые ей выходные данные. Среда MATLAB имеет отдельную форму редактора программных файлов (скриптов, функций, классов) – Editor, которая показана на Рис.4.

---

[1] Описание рабочей среды производится на примере версии MATLAB 7.10.0 (R2010a) для Windows. Несмотря на некоторые отличия в интерфейсе различных версий программной среды MATLAB, набор команд для проведения численных расчетов, требуемых в лабораторных работах данного курса, остался практически неизменным.

Запуск отдельных команд производится путём их ввода в «Command Window» (II), вычисленные в результате работы скрипта или отдельных команд переменные отображаются в панели «Workspace» (III), а история запуска операций в окне «Command History» (IV).

Альтернативой запуска отдельных команд является создание файла скрипта (*.m) путём выбора пункта меню «File» -> «New» -> «Script», который также можно запускать в режиме отладки в окне редактора Editor (Рис.4).

Ниже приведены основные сведения, требуемые при работе с MATLAB.

Пример 1. Решение системы линейных уравнений

A=[1 0.5 1; 0.5 2 0.5; 1 0.5 3]   c=[1 2 3]'   *// ввод начальных данных*

*b=inv(A)*c*   *// решение системы*

$$\mathrm{b} = \begin{bmatrix} -0.4286 \\ 0.8571 \\ 1 \end{bmatrix}$$ *// результат*

Пример 2. Вычисление суммы ряда

syms k

r = symsum(1/sym('k!'), k, 0, 10)

r = 2.7183

Пример 3. Вычисление интеграла

syms x

r = int(1/(1+x^2), 0, 1)

r = 0.7854

Пример 4. Решение алгебраического уравнения второй степени в аналитическом виде

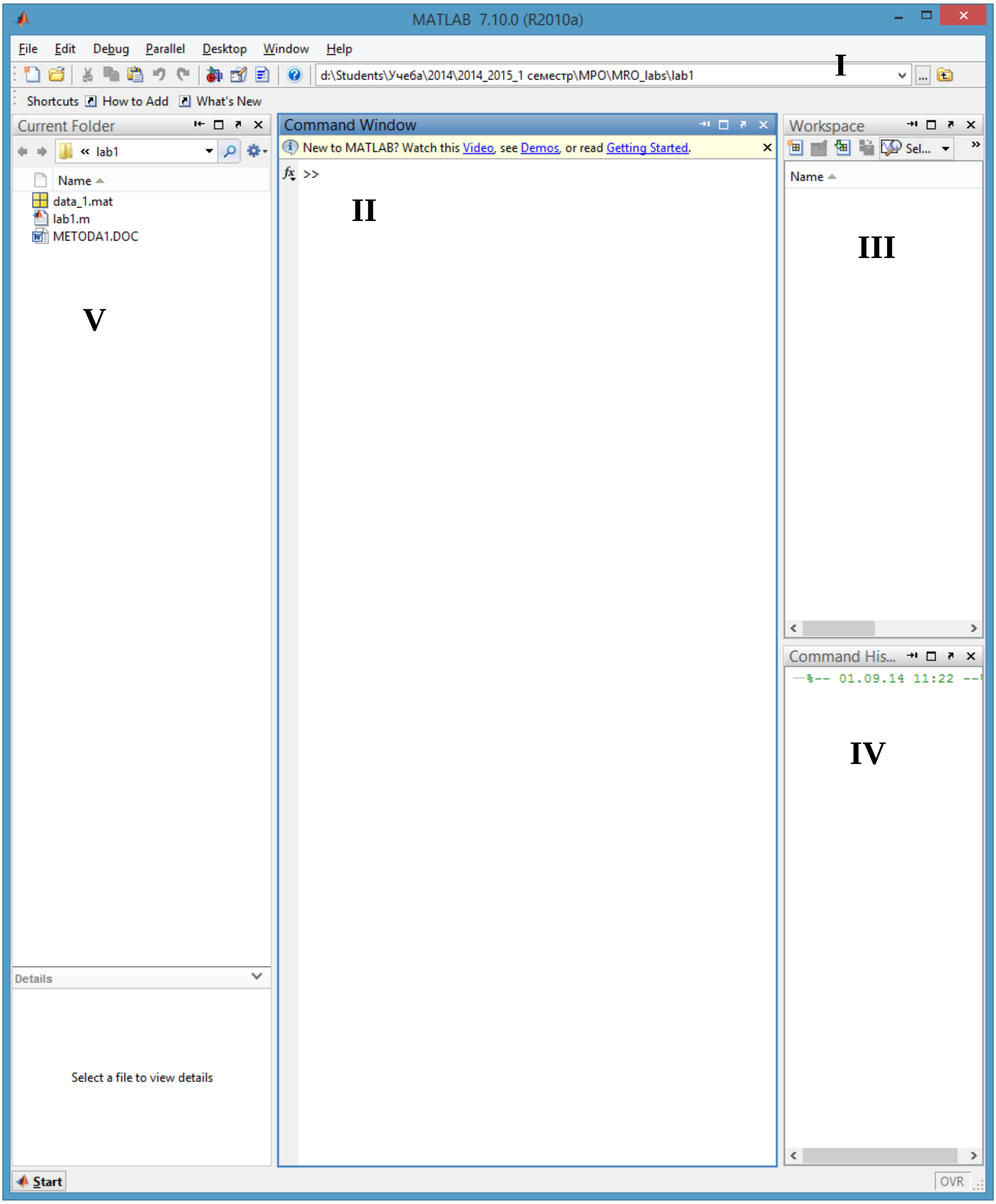
syms x

*r=solve(x^2+5*x-25)*

*r= -(5*5^(1/2))/2 - 5/2*

*(5*5^(1/2))/2 - 5/2*

Рис.2 *Примеры математического программирования на MATLAB*



I –панель инструментов, II – панель ввода команд,
III – панель рабочего пространства, IV – история команд, V – рабочий каталог,
Рис.3 *Рабочая область MATLAB и элементы управления*

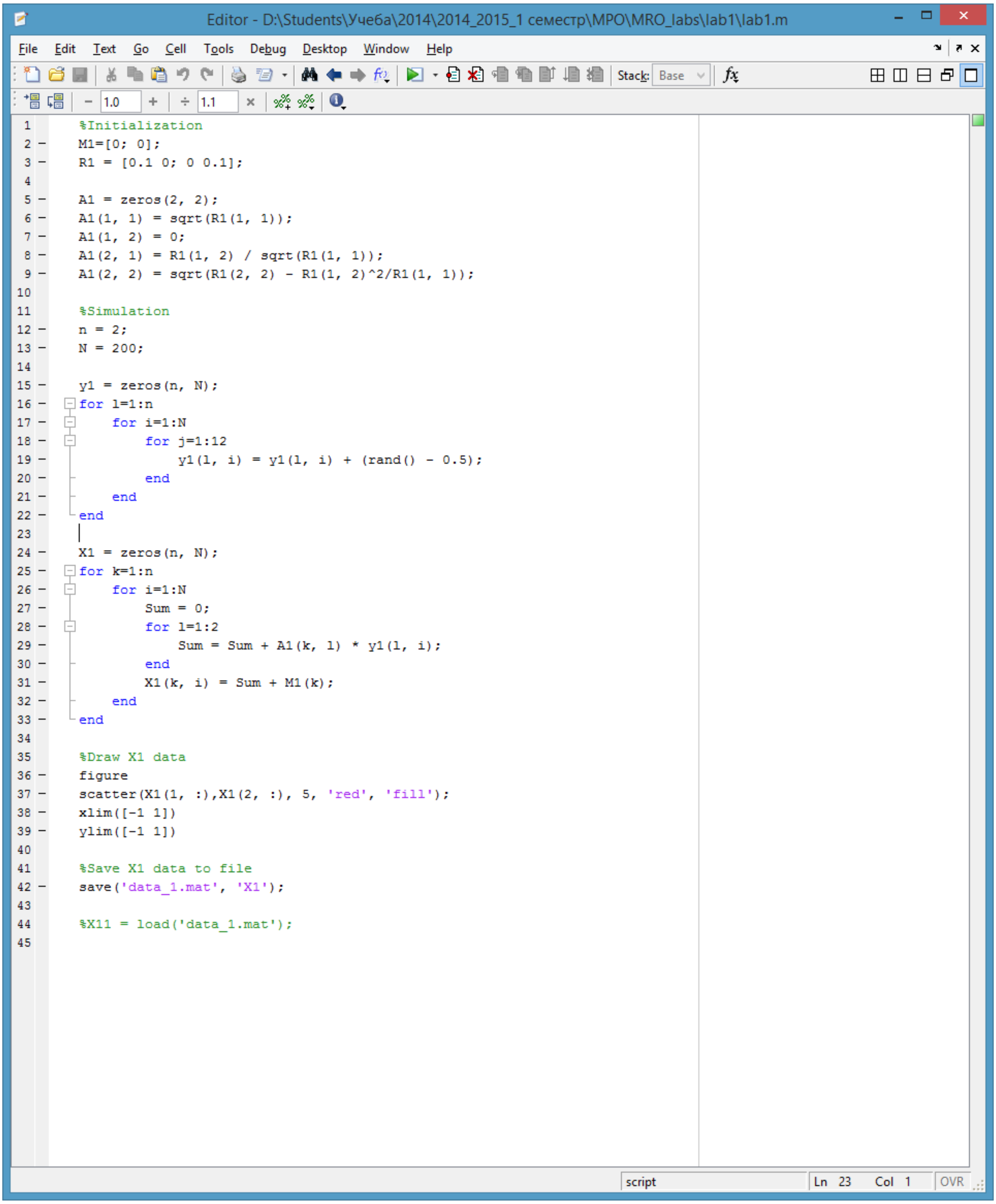


Рис.4 Среда редактирования программных файлов MATLAB

2.2. Построение математических выражений

Выражения в MATLAB могут иметь одну из следующих форм:

выражение… - производит вычисление значения выражения;

переменная = *выражение* - задание выражения для вычисления переменной;

переменная := n1:step:n2 - задание пределов изменения переменной;

[вых1, вых2,…]=функция(арг1, арг2,...) - вызов функции со списком аргументов «*арг1*», «*арг2*» и т.д., возвращающей значения переменных «вых1», «вых2» и т.д.; функция хранится в отдельном файле, который можно создать путем вызова команды «File» -> «New» -> «Function»

При наборе выражений используют следующие команды редактора MATLAB.

оператор	**ввод на экране**
пределы изменения переменной (от x до z с шагом y)	x:y:z
скобки	(x)
факториал	factorial(x)
степень	x^y
корень квадратный	sqrt(x)
модуль, детерминант матрицы	abs(A), det(A)
сравнения	x<y или x>y x>=y или x<=y
не равно	x~=y
индекс	x(i)
двойной индекс	M(i,j)
транспонированная матрица	M’
сумма элементов вектора	sum(x)
произведение элементов вектора	prod(x)

2.3. Обзор встроенных функции

sin(z), cos(z),

tan(z) - тригонометрические функции (аргумент в радианах);

asin(z), acos(z),

atan(z) - обратные тригонометрические функции (результат в радианах);

sinh(z), asinh(z),

cosh(z), acosh(z),

tanh(z), atanh(z) - гиперболические функции;

exp(z) - e^z;

log(z), log10(z)	- функции логарифма числа *z;*
erf(x)	- функция ошибок $\int_0^x \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt$;
gamma(x)	- гамма-функция Эйлера ($-3 \le x \le 3$);
randi(x)	- датчик случайных чисел, равномерно распределенных от 0 до *x*;
while(условие) *[набор команд]* *end*	- цикл с постусловием;
if (условие) *[набор команд]* *end*	- условный оператор;
mean(v)	- среднее значение вектора *v*;
var(v)	- дисперсия вектора *v*;
cov(A)	- ковариационная матрица матрицы A;
std(v)	- среднеквадратическое отклонение вектора *v*.

2.4. Прикладная программа в MATLAB

Структура скрипта или функции в среде MATLAB напоминает структуру обычной прикладной программы на традиционном языке программирования: она, как правило, содержит блок инициализации, блок расчетов и блок отображения результатов, каждый из которых не является обязательным.

В качестве примера программы MATLAB ниже приведена программа моделирования *N* значений двумерного нормально распределенного случайного вектора.

Текст программы в MATLAB	**Комментарии**
`M1 = [0; 0];` `R1 = [0.1 0; 0 0.1];`	Задание параметров нормального закона распределения

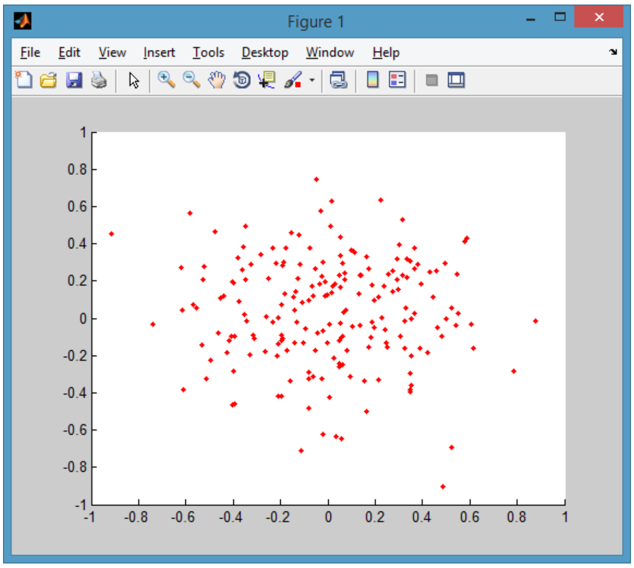
`A1 = zeros(2, 2);`
`A1(1, 1) = sqrt(R1(1, 1));`
`A1(1, 2) = 0;`
`A1(2, 1) = R1(1, 2) / sqrt(R1(1, 1));`
`A1(2, 2) = sqrt(R1(2, 2) -`
`                R1(1, 2)^2/R1(1, 1));`
	Определение параметров линейного преобразования

<pre>A1 =
    0.3162   0
         0   0.3162</pre>
	Отображение полученного результата для матрицы линейного преобразования

<pre>n = 2;
N = 200;
y1 = zeros(n, N);
X1 = zeros(n, N);</pre>
	Вспомогательные переменные, отвечающие за двухкомпонентность вектора, число выборочных значений, и за процесс генерации стандартной нормально распределенной случайной величины.

<pre>for l=1:n
    for i=1:N
        for j=1:12
            y1(l, i) = y1(l, i) +
(rand() - 0.5);
        end
    end
end</pre>
	Генерация *N* реализаций случайного вектора, компоненты которого – суть независимые и нормально распределенные $N(0,1)$ случайные величины.

<pre>for k=1:n
    for i=1:N
        Sum = 0;
        for l=1:2
            Sum = Sum + A1(k, l) *
y1(l, i);
        end
        X1(k, i) = Sum + M1(k);
    end
end</pre>
	Генерация *N* реализаций случайного вектора с требуемым нормальным законом распределения $N(\overline{M}, B)$.

<pre>figure
scatter(X1(1, :),X1(2, :), 5, 'red',
'fill');
xlim([-1 1])
ylim([-1 1])</pre> 
	Графическое отображение результатов моделирования нормально распределенного случайного вектора в пределах по осям *x* и *y* [-1, 1]. Каждая точка отображается закрашенной («fill») красным («red») цветом окружностью с радиусом «5».

<pre>save('data_1.mat', 'X1');</pre>	Сохранение значений переменной в .mat файл

3. ЛИТЕРАТУРА

1. Анисимов Б.В., Курганов В.Д., Злобин В.К. Распознавание и цифровая обработка изображений. - М.: Высшая школа, 1983. - 295 с.
2. Дуда Р., Харт П. Распознавание образов и анализ сцен: Пер. с англ. - М.: Мир, 1976. - 512 с.
3. Ту Дж., Гонсалес Р. Принципы распознавания образов: Пер. с англ. - М.: Мир, 1978. - 412с.
4. Фукунага К. Введение в статистическую теорию распознавания образов: Пер. с англ. - М.: Наука, 1979. - 368с.
5. Курбатова Е.А. MATLAB 7. Самоучитель. Издательство: Вильямс. Год издания: 2005г. 256 стр.
6. Lynch, Stephen. Dynamical Systems with Applications using MATLAB, 2004.

4. ПОРЯДОК ВЫПОЛНЕНИЯ ЛАБОРАТОРНОЙ РАБОТЫ

4.1. Исходные данные

- Вариант задания (предоставляется преподавателем);
- математические ожидания для трех наборов двумерных нормально распределенных случайных векторов (из соответствующего варианта задания);
- два бинарных вектора;
- исполняемый в системе MATLAB файл, необходимый для выполнения лабораторной работы: lab1.m (предоставляется преподавателем).

4.2. Общий план выполнения работы

1. Разработать алгоритм моделирования нормально распределенного случайного вектора с заданными математическим ожиданием и корреляционной матрицей.
2. Смоделировать и изобразить графически обучающие выборки объема $N=200$ для двух нормально распределенных двумерных случайных векторов с заданными математическими ожиданиями и самостоятельно подобранными равными корреляционными матрицами.
3. Смоделировать и изобразить графически обучающие выборки объема $N=200$ для трех нормально распределенных двумерных случайных векторов с заданными

математическими ожиданиями и с неравными корреляционными матрицами, которые выбрать самостоятельно.

4. На основании полученных выборок найти точечные оценки параметров нормального закона для каждого из распределений.
5. Смоделировать обучающие выборки объема N=200 двух бинарных случайных векторов с распределениями, которые обеспечивают вероятность изменения указанной в представителе компоненты случайного вектора равную $p = 0.3$.

4.3. Содержание отчета

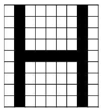
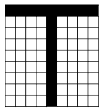
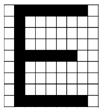
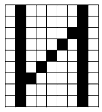
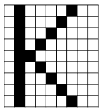
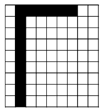
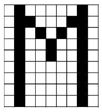
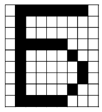
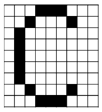
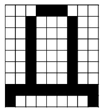
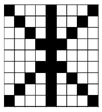
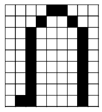
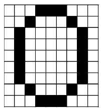
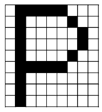
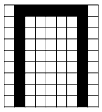
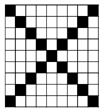
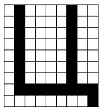
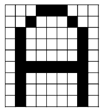
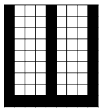
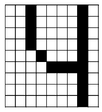
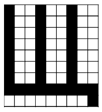
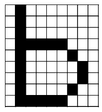
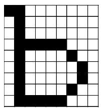
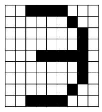
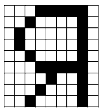
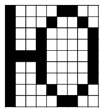
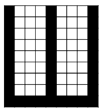
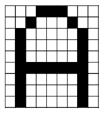
Отчет по работе должен содержать:

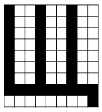
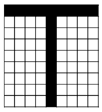
- исходные параметры моделируемых нормальных распределений; их оценки, полученные по обучающим выборкам, расстояния Бхатачария и Махаланобиса;
- графическое изображение значений векторов, полученных в п.2 и п.3, и имена файлов (с расширением .mat), в которые они записаны;
- распределения бинарных случайных векторов и имена записанных файлов, содержащих их реализации (с расширением .mat).

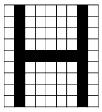
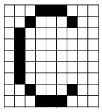
5. КОНТРОЛЬНЫЕ ВОПРОСЫ

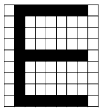
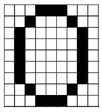
1. Алгоритм моделирования нормально распределенного случайного вектора.
2. Вид матрицы линейного преобразования, используемой для моделирования нормально распределенного случайного вектора.
3. Оценивание параметров нормального закона распределения.
4. Выражения для рекуррентного оценивания параметров нормального закона распределения.
5. Меры близости нормальных распределений.
6. Инвариантность расстояний к линейным преобразованиям.
7. Характер линейного преобразования, обеспечивающего инвариантность евклидового расстояния.
8. Алгоритм моделирования бинарного случайного вектора с независимыми координатами.
9. Отличие среды математического программирования MATLAB от традиционных языков программирования.
10. Структура прикладной программы в MATLAB.

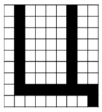
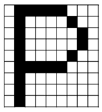
6. ВАРИАНТЫ ЗАДАНИЙ

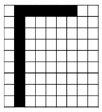
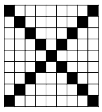
Вариант	Математические ожидания трех наборов нормально распределенных случайных векторов	Представители бинарных случайных векторов, □ ~ "0", ■ ~ "1"
1.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$	 
2.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$	 
3.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	 
4.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	 
5.	$\overline{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$	 
6.	$\overline{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$	 
7.	$\overline{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	 
8.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$	 
9.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -2 \\ 1 \end{pmatrix}.$	 
10.	$\overline{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}.$	 
11.	$\overline{M}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$	 
12.	$\overline{M}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$	 
13.	$\overline{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}.$	 
14.	$\overline{M}_1 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}.$	 

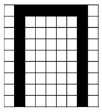
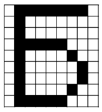
15. $\overline{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$  

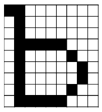
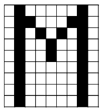
16. $\overline{M}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$  

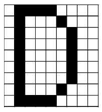
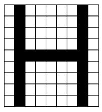
17. $\overline{M}_1 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$  

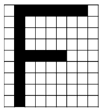
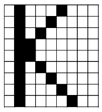
18. $\overline{M}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 2 \\ 2 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$  

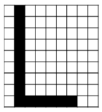
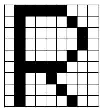
19. $\overline{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 2 \\ -2 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$  

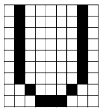
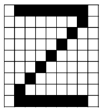
20. $\overline{M}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -2 \\ -2 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$  

21. $\overline{M}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -2 \\ 2 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$  

22. $\overline{M}_1 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}.$  

23. $\overline{M}_1 = \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}.$  

24. $\overline{M}_1 = \begin{pmatrix} -2 \\ 0 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}.$  

25. $\overline{M}_1 = \begin{pmatrix} 0 \\ -2 \end{pmatrix}, \quad \overline{M}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \overline{M}_3 = \begin{pmatrix} 2 \\ 0 \end{pmatrix}.$  

СОДЕРЖАНИЕ

Учебное издание

МОДЕЛИРОВАНИЕ ИСХОДНЫХ ДАННЫХ ДЛЯ РЕШЕНИЯ ЗАДАЧ
РАСПОЗНАВАНИЯ ОБРАЗОВ

Методические указания к лабораторной работе № 1
по курсу «Методы распознавания образов»

Составители: Коломиец Эдуард Иванович
Мясников Владислав Валерьевич

Самарский государственный аэрокосмический университет
имени академика С.П. Королева
443086, Самара, Московское шоссе, 34

Отпечатано на кафедре «Техническая кибернетика»
Тираж 20 экз.