Part VI

Mahalanobis–Taguchi System (MTS)

20 Mahalanobis–Taguchi System

20.1. Introduction

Many people remember the RCA Victor Company, with its logo of a dog listening to "his master's voice." Dogs can easily recognize their masters' voices, but it is not easy for a machine to be created that recognizes its owner by voice recognition. Painting lovers can easily recognize the impressionist paintings of Vincent van Gogh. Can we develop a machine that will recognize van Gogh painting? Classical music lovers can distinguish the music composed by Johann Sebastian Bach from that of others. But even using the most advanced computers, it is not easy for human-made machines to recognize human voice, face, handwriting, or printing.

Human beings have the ability to recognize visual patterns because we have some 9 million optical sensing cells in our eyes. Assuming that the time to treat a captured image is 0.04 second, we can treat 225 million pieces of information in 1 second. We do not necessarily try to memorize such a vast amount of information, but our sensing cells are efficient and simplify and screen out unnecessary pieces of information. If we could develop a system that made it simple enough to use current personal computers for pattern recognition, there would be numerous applications in various areas.

20.2. What Is the MTS?

MTS is the acronym for the *Mahalanobis–Taguchi system*. P. C. Mahalanobis was a famous Indian statistician who established the Indian Statistical Institute. In 1930, Mahalanobis introduced a statistical tool called the *Mahalanobis distance* (MD), used to distinguish the pattern of a certain group from other groups. Mahalanobis used the Mahalanobis distance for an archaeological application to classify excavated bones and to make judgments as to which dinosaur a variety of sample bones belonged to. Another application was ethnological, to characterize differences among Asian races and tribes. The main objective of his application was to make statistical judgments to distinguish one group from another.

The Mahalanobis–Taguchi system is Taguchi's design of a systematic method for using the Mahalanobis distance. The objective of MTS is to develop and optimize a diagnostic system with a measurement scale of *abnormality*. In MTS, the signal-to-noise (SN) ratio is used to assess the effectiveness of a system. Moreover, Taguchi used the system not only for diagnosis but also for forecasting/prediction systems. Thus, MTS is used to develop and optimize a system of multivariable diagnosis, pattern recognition, and prediction of occurrence of particular events.

20.3. Challenges Appropriate for the MTS

Diagnosis or pattern recognition is not easy, since it is based on multivariable data consisting of both continuous and attribute data, with correlations among those variables. For example:

1. Physicians conduct diagnoses on patients; observe patients for symptoms, and look at patient data such as blood pressure, pulse, blood test results, and past data on the patient. A good doctor will come to a valid diagnostic result, and a poor one would misdiagnose the problem. Misdiagnosis can be very costly.

2. An inspection system must make a judgment as to whether a product or system can pass or fail based on multiple specifications. The same can be said for monitoring a complex manufacturing process.

3. It is difficult to evaluate and improve discrimination power among handwritten letters. There is no perfect handwriting, voice, fingerprint, or facial expression recognition system.

4. It is extremely difficult to predict earthquakes. How can we predict when and where a big one will occur?

To illustrate the difficulty of any of these, let's take a very simple example with two variables, x_1 and x_2. Let x_1 be the weight of an American male and x_2 be the height of an American male. Suppose that a weight is 310 pounds. If you look at this data, you may think that this person (John) is very heavy but not necessarily abnormal, since there are many healthy American males who weigh around 310 (Figure 20.1).

Suppose that John's height is 5 feet. If you look at his height by itself, you may think John is a very short person but not that there is necessarily anything abnormal about John, since there are many healthy American males whose height is around 5 feet (Figure 20.2).

On the other hand, if you look at both pieces of data on John, $x_1 = 310$ pounds and $x_2 = 5.0$ feet, you would think something was definitely out of the ordinary about John. You would think that he is not necessarily normal. This is because, as everyone knows, there is a correlation between weight and height. Suppose that we can define the "normal" group in terms of weight and height (Figure 20.3). That would have the shape of an ellipse because of the correlation between weight and height. We can clearly see that John is outside the normal group. The distance between the center of a normal group and a person under diagnosis is the Mahalanobis distance.

From this discussion, you can see the following points:

1. For an inspection system, two characteristics meeting the specification do not guarantee that it is a good product because of correlation (interaction).

2. It is easy to see the pattern when the number of variables, k, is equal to 2. What if $k = 5$, $k = 10$, $k = 20$, $k = 40$, $k = 120$, $k = 400$, or $k = 1000$? The greater the number of variables, the more difficult it is to see the pattern.

The steps to design and optimize the MTS are as follows:

❏ *Task 1:* Generation of normal space

    Step 1. Define the normal group.

    Step 2. Define k variables ($1 < k < 1000$).

    Step 3. Gather data from the normal group on k variables with sample size n, $n \gg k$.

    Step 4. Calculate the MD for each sample from the normal group.

❏ *Task 2:* Confirmation of discrimination power

    Step 5. Gather data on k variables for samples from outside the normal group; sample size $= r$.
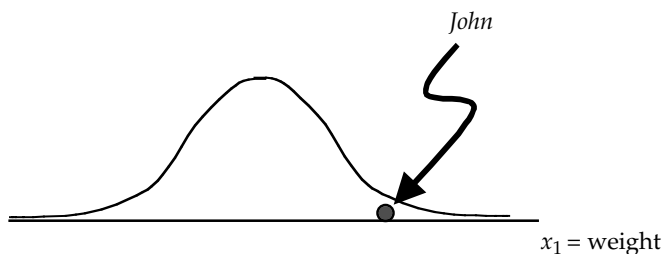
Figure 20.1
John's weight

John

x_2 = height

Figure 20.2
John's height

Step 6. Calculate the MD for each sample from outside the normal group.

Step 7. Evaluate the discrimination power.

❏ *Task 3:* Identification of critical variables and optimization

Step 8. Optimize the MTS. Evaluate critical variables. Optimize the number of variables. Define the optimum system.

There are two primary methods of computing the MD. One is to use an inverted matrix and the other is to use Gram–Schmidt orthogonal vectors, which we refer to as MTGS. First, a simple case study with inverted matrix method is illustrated. MTGS will be introduced later.

x_2 = height

*Normal Space
(Mahalanobis Space)
(Reference Space)*

MD = Mahalanobis distance

John

Figure 20.3
Both height and weight

x_1 = weight
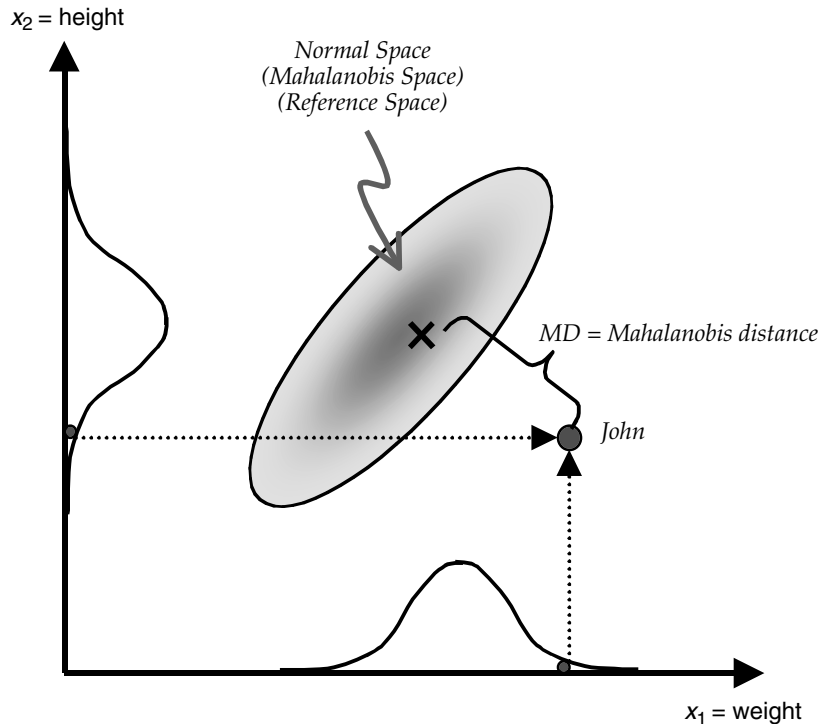
20.4. MTS Case Study: Liver Disease Diagnosis*

Step 1. Define the normal group: people with a healthy liver $(n = 200)$. Physicians specializing in liver disease identified 200 people who had healthy livers. They defined the normal group as any person with a healthy liver.

Step 2. Define the variables $(k = 15)$. The following variables were selected:

x_1	Age	x_6	LDH	x_{11}	TG
x_2	Gender	x_7	ALP	x_{12}	PL
x_3	TP	x_8	GTP	x_{13}	C_r
x_4	Alb	x_9	LAP	x_{14}	BUN
x_5	ChE	x_{10}	Tch	x_{15}	UA

Physicians identified these 15 variables, 13 from blood test results. Variables can be continuous or discrete. In this case, gender was a discrete variable.

Step 3. Generate a database of a normal group for the variables selected. For a discrete variable with two classes, such as gender, simply assign a number for each class: say, 1 for male and 10 for female (Table 20.1).

Step 4a. Compute the mean and standard deviation for each variable and normalize the data. See Table 20.2.

$$Z = \frac{x - \bar{x}}{\sigma_x}$$

Step 4b. Compute the correlation matrix. See Table 20.3.

Step 4c. Compute the inverse matrix of the correlation matrix. See Table 20.4.

Step 4d. Compute MD for each sample from the normal group. Each person in the normal group is a vector of $(z_1, \ldots, z_{15})$. For each person, MD is given by the following equation and by Table 20.5 and Figure 20.4:

$$\mathrm{MD} = \frac{1}{k} ZA^{-1}Z^T$$

$$= \frac{1}{k} (z_1, z_2, \ldots, z_{15}) \ (A^{-1}) \begin{pmatrix} z_1 \\ \vdots \\ z_{15} \end{pmatrix}$$

The distribution of Mahalanobis distances for samples from the normal group typically look like Figure 20.5, below, shown in Figure 20.6 as a curve. The average is 1.00. Note that it is very rare to have an MD value greater then 5.0; an MD for an abnormal sample should become much larger than 5.0; MD provides a measurement scale for abnormality.

*This case study is presented by courtesy of T. Kanetaka, Teishin Hospital, Tokyo.

Table 20.1
Data for normal group

Sample	Age x_1	Gender x_2	TP x_3	Alb x_4	ChE x_5	LDH x_6	ALP x_7	GTP x_8	LAP x_9	Tch x_{10}	TG x_{11}	PL x_{12}	Cr x_{13}	BUN x_{14}	UA x_{15}
1	59	10	7	4.2	0.7	190	7.7	12	275	220	140	228	0.8	13	3.8
2	51	1	7	4.4	0.98	217	6.3	23	340	225	87	227	1.1	15	4.6
3	53	1	6.7	4.2	0.83	220	9.2	16	278	220	67	222	0.8	17	3.2
4	52	1	6.6	4.3	0.9	204	4.6	15	289	171	59	175	0.8	13	3.8
5	52	1	6.7	3.9	0.97	198	5.2	13	312	192	51	203	1	14	4.1
6	56	1	7.2	4	0.93	188	6.1	16	304	216	86	213	1	13	4.2
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
198	35	10	7.3	4.3	0.7	203	6.1	19	343	165	59	178	1.3	14	6.4
199	35	10	7.1	4.4	0.86	190	4	36	358	190	155	200	1.4	19	5.7
200	53	1	6.8	4	0.88	225	8.2	18	330	189	124	203	1	17	4.3
Avg.	39.7	6.9	7.1	4.3	0.9	195.4	5.5	22.6	332.1	184.0	84.7	192.1	1.2	15.1	5.0
Sum	10.34	4.30	0.30	0.14	0.10	22.93	1.36	12.92	55.76	22.21	34.86	20.57	0.19	2.88	1.22

Table 20.2
Mean and standard deviation for normal group

Sample	Age z_1	Gender z_2	TP z_3	Alb z_4	ChE z_5	LDH z_6	ALP z_7	GTP z_8	LAP z_9	Tch z_{10}	TG z_{11}	PL z_{12}	Cr z_{13}	BUN z_{14}	UA z_{15}
1	1.87	0.73	-0.19	-0.66	-1.91	-0.24	1.59	-0.82	-1.02	1.62	1.59	1.75	-1.9	-0.74	-1.01
2	1.09	-1.36	-0.19	0.77	1	0.94	0.56	0.03	0.14	1.85	0.06	1.7	-0.33	-0.04	-0.35
3	1.29	-1.36	-1.2	-0.66	-0.56	1.07	2.7	-0.51	-0.97	1.62	-0.51	1.46	-1.9	0.65	-1.5
4	1.19	-1.36	-1.54	0.06	0.17	0.37	-0.69	-0.59	-0.77	-0.58	-0.74	-0.83	-1.9	-0.74	-1.01
5	1.19	-1.36	-1.2	-2.81	0.9	0.11	-0.25	-0.74	-0.36	0.36	-0.97	0.53	-0.85	-0.39	-0.76
6	1.58	-1.36	0.49	-2.09	0.48	-0.32	0.41	-0.51	-0.5	1.44	0.04	1.02	-0.85	-0.74	-0.68
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
198	-0.45	0.73	0.83	0.06	-1.91	0.33	0.41	-0.28	0.2	-0.85	-0.74	-0.68	0.73	-0.39	1.12
199	-0.45	0.73	0.15	0.77	-0.24	-0.24	-1.13	1.04	0.46	0.27	2.02	0.39	1.25	1.35	0.56
200	1.29	-1.36	-0.86	-2.09	-0.04	1.29	1.96	-0.35	-0.04	0.23	1.13	0.53	-0.85	0.65	-0.6
Avg.	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Sum	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00

Table 20.3
Correlation matrix

	z_1	z_2	z_3	z_4	z_5	z_6	z_7	z_8	z_9	z_{10}	z_{11}	z_{12}	z_{13}	z_{14}	z_{15}
z_1	**1.0000**	−0.2968	−0.2993	−0.4168	−0.0570	0.2084	0.3335	0.0234	−0.1119	0.2343	0.1405	0.2658	−0.2305	−0.0292	−0.2108
z_2	−0.2968	**1.0000**	0.0811	0.3874	0.1996	−0.1084	−0.0581	0.3638	0.3934	−0.1966	0.2707	−0.2189	0.6490	0.2675	0.5616
z_3	−0.2993	0.0811	**1.0000**	0.3966	0.0763	0.0454	−0.0149	0.1588	0.1762	0.0785	0.1225	0.0845	0.1376	−0.1167	0.2022
z_4	−0.4168	0.3874	0.3966	**1.0000**	0.1274	0.0139	−0.1341	0.1769	0.2266	−0.0864	0.0423	−0.0943	0.2767	0.0961	0.3684
z_5	−0.0570	0.1996	0.0763	0.1274	**1.0000**	0.1472	0.0264	0.3305	0.4005	0.2891	0.3198	0.2866	0.2535	0.0293	0.3500
z_6	0.2084	−0.1084	0.0454	0.0139	0.1472	**1.0000**	0.2230	0.2195	0.2285	0.2137	0.1051	0.2488	0.0666	0.0919	0.0514
z_7	0.3335	−0.0581	−0.0149	−0.1341	0.0264	0.2230	**1.0000**	0.1754	0.1615	0.1806	0.1523	0.1783	−0.0348	−0.0869	−0.0430
z_8	0.0234	0.3638	0.1588	0.1769	0.3305	0.2195	0.1754	**1.0000**	0.7297	0.2057	0.4462	0.2110	0.3851	0.0534	0.4663
z_9	−0.1119	0.3934	0.1762	0.2266	0.4005	0.2285	0.1615	0.7297	**1.0000**	0.1855	0.3378	0.1634	0.4610	0.0514	0.4427
z_{10}	0.2343	−0.1966	0.0785	−0.0864	0.2891	0.2137	0.1806	0.2057	0.1855	**1.0000**	0.3036	0.9334	0.0381	0.0173	0.1354
z_{11}	0.1405	0.2707	0.1225	0.0423	0.3198	0.1051	0.1523	0.4462	0.3378	0.3036	**1.0000**	0.3014	0.2553	−0.0288	0.3791
z_{12}	0.2658	−0.2189	0.0845	−0.0943	0.2866	0.2488	0.1783	0.2110	0.1634	0.9334	0.3014	**1.0000**	0.0204	0.0349	0.1191
z_{13}	−0.2305	0.6490	0.1376	0.2767	0.2535	0.0666	−0.0348	0.3851	0.4610	0.0381	0.2553	0.0204	**1.0000**	0.2485	0.5713
z_{14}	−0.0292	0.2675	−0.1167	0.0961	0.0293	0.0919	−0.0869	0.0534	0.0514	0.0173	−0.0288	0.0349	0.2485	**1.0000**	0.1447
z_{15}	−0.2108	0.5616	0.2022	0.3684	0.3500	0.0514	−0.0430	0.4663	0.4427	0.1354	0.3791	0.1191	0.5713	0.1447	**1.0000**

Table 20.4
Inverse matrix

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	1.6383	0.1542	0.3455	0.3415	0.0986	-0.2063	-0.3587	-0.2414	0.2598	0.1133	-0.2578	-0.3422	0.1152	-0.0397	0.0843
2	0.1542	2.6720	0.2274	-0.4176	-0.0459	0.3648	0.1303	-0.1838	-0.1639	0.2613	-0.3976	0.4678	-1.0153	-0.3740	-0.5164
3	0.3455	0.2274	1.3519	-0.4626	0.0908	-0.0352	-0.0652	-0.0759	-0.0251	0.0999	-0.1414	-0.2343	-0.0860	0.1833	-0.0536
4	0.3415	-0.4176	-0.4626	1.5731	-0.0202	-0.1313	0.0640	0.0007	-0.0395	0.0292	0.1347	0.0027	0.1499	-0.0570	-0.2785
5	0.0986	-0.0459	0.0908	-0.0202	1.3626	-0.0757	0.0720	0.0646	-0.3733	-0.0687	-0.1840	-0.2407	0.0164	0.0470	-0.2076
6	-0.2063	0.3648	-0.0352	-0.1313	-0.0757	1.2503	-0.1651	-0.0903	-0.2222	0.2280	0.0225	-0.2908	-0.1468	-0.1634	0.0055
7	-0.3587	0.1303	-0.0652	0.0640	0.0720	-0.1651	1.2354	-0.0508	-0.1969	-0.1598	-0.0394	0.0640	0.0628	0.1115	0.1114
8	-0.2414	-0.1838	-0.0759	0.0007	0.0646	-0.0903	-0.0508	2.5327	-1.5365	0.3081	-0.3927	-0.3686	0.0939	0.0036	-0.3568
9	0.2598	-0.1639	-0.0251	-0.0395	-0.3733	-0.2222	-0.1969	-1.5365	2.6016	-0.4256	0.0977	0.3556	-0.3981	0.1034	0.0556
10	0.1133	0.2613	0.0999	0.0292	-0.0687	0.2280	-0.1598	0.3081	-0.4256	8.0089	-0.1736	-7.3840	-0.1133	0.0849	-0.1908
11	-0.2578	-0.3976	-0.1414	0.1347	-0.1840	0.0225	-0.0394	-0.3927	0.0977	-0.1736	1.5448	-0.1540	0.0449	0.1621	-0.2435
12	-0.3422	0.4678	-0.2343	0.0027	-0.2407	-0.2908	0.0640	-0.3686	0.3556	-7.3840	-0.1540	8.3166	-0.0523	-0.2711	-0.0348
13	0.1152	-1.0153	-0.0860	0.1499	0.0164	-0.1468	0.0628	0.0939	-0.3981	-0.1133	0.0449	-0.0523	2.1288	-0.1681	-0.4937
14	-0.0397	-0.3740	0.1833	-0.0570	0.0470	-0.1634	0.1115	0.0036	0.1034	0.0849	0.1621	-0.2711	-0.1681	1.1955	0.0172
15	0.0843	-0.5164	-0.0536	-0.2785	-0.2076	0.0055	0.1114	-0.3568	0.0556	-0.1908	-0.2435	-0.0348	-0.4937	0.0172	2.0420

Table 20.5
Mahalanobis distance for 200 samples from healthy group

|																			
2.11	0.61	1.13	0.75	0.82	0.74	0.57	0.86	0.64	1.20	0.85	0.69	0.66	0.99	0.76	0.52	0.81	0.77	1.29	1.16
0.58	0.95	0.64	1.15	0.67	0.72	0.83	0.89	0.69	1.22	0.64	0.88	1.04	1.19	1.25	0.72	1.09	0.47	2.77	1.55
1.41	1.11	0.43	0.66	2.63	0.74	1.12	1.04	0.93	1.25	1.52	0.69	1.69	0.72	0.72	0.41	1.81	1.06	0.95	0.59
0.70	1.02	0.90	0.76	0.82	1.00	0.90	0.56	2.59	0.70	0.90	1.09	0.52	0.88	1.32	0.63	1.64	0.98	1.39	0.53
0.95	0.86	1.18	0.84	0.50	0.92	0.96	0.94	0.82	0.91	1.75	0.37	0.65	0.46	0.81	0.45	2.76	1.47	0.49	0.76
0.82	0.58	1.07	0.97	3.34	0.84	0.61	1.65	0.82	1.88	0.96	1.13	1.58	0.78	1.63	0.94	0.73	0.51	0.64	0.88
1.99	0.78	0.91	1.26	2.10	1.35	0.79	0.51	0.72	0.97	1.04	0.59	0.61	0.79	1.45	0.83	2.21	0.69	0.66	1.19
1.06	0.97	0.80	0.62	1.08	0.61	0.61	0.81	0.79	0.52	0.83	0.49	0.68	1.62	1.03	0.98	0.88	1.11	0.80	0.64
1.35	0.78	1.67	0.61	0.95	0.67	0.86	1.08	0.78	0.90	0.88	0.71	0.74	1.37	0.88	0.97	1.24	0.80	0.44	1.18
1.70	1.30	0.45	1.46	2.05	1.20	1.25	0.88	1.20	0.71	0.57	0.55	1.87	1.20	1.03	0.69	0.90	0.77	0.75	0.91
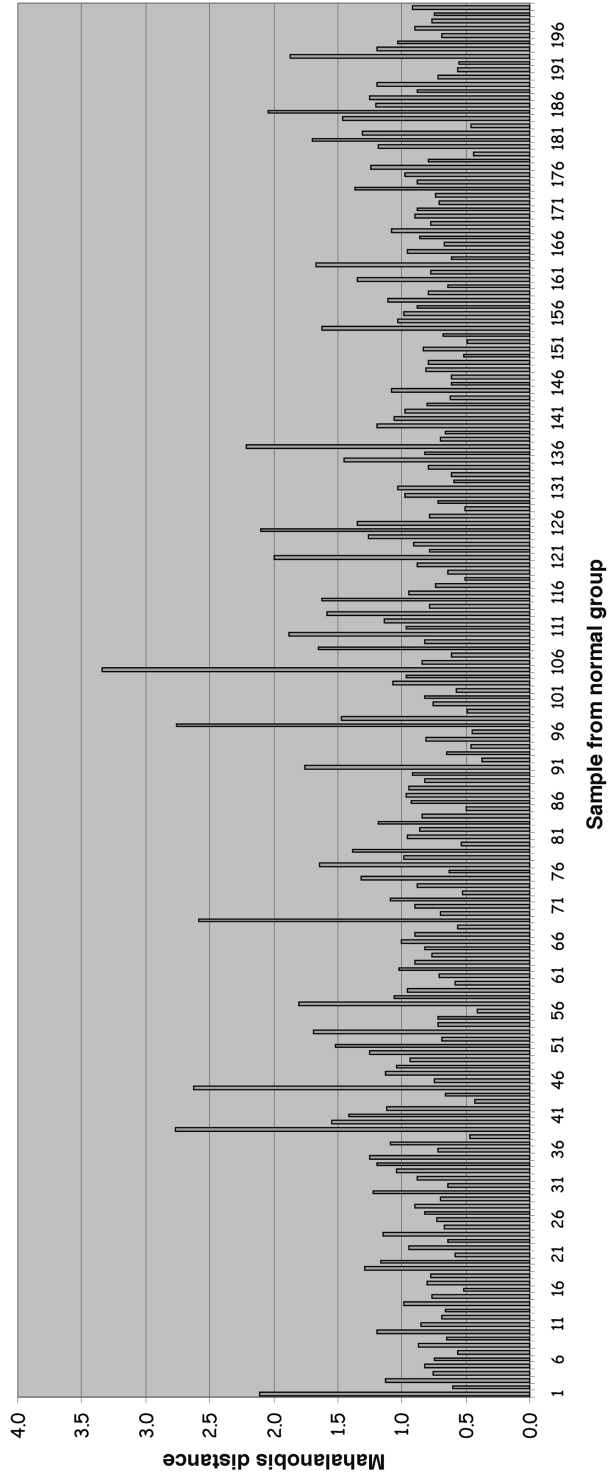
Figure 20.4
Sample showing Table 20.5 data

Figure 20.5
MD distribution
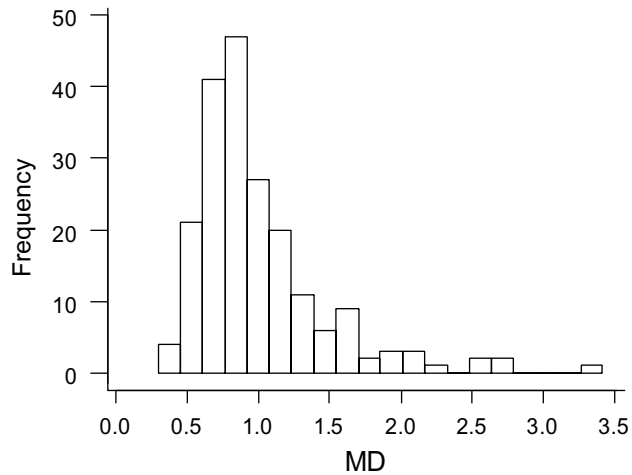
At this point it is important to recognize the following:

❏ The objective of MTS is to develop a measurement scale that measures abnormality. What we have just done is to define the zero point and one unit.

❏ MTS measures abnormality. Only the normal group has a population.

Now we will assess the discrimination power by gathering information.

Step 5. Gather data on k variables for samples from outside the normal group. See Table 20.6.

Step 6. Calculate the MD for each sample from outside the normal group. See Table 20.7.

Step 7. Evaluate The discrimination power. See Figure 20.7. Just by eyeballing Figure 20.7, one can see that the discrimination power is very good. In some studies it may be sufficient to have an MTS that can discriminate *x* percent of bad samples. For example, there is a situation called "no trouble found" for those samples that come back for warranty claims. In that situation, discriminating only 25% can be extremely beneficial.

After confirming a good discrimination, we are ready to optimize the MTS.

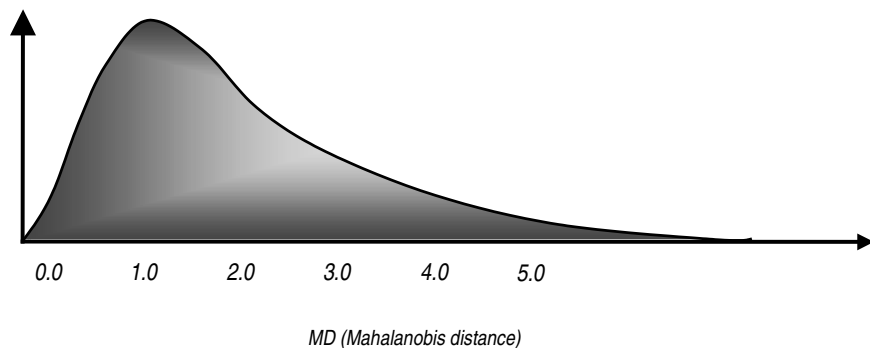**Figure 20.6**
Distribution for Figure
20.6

MD (Mahalanobis distance)

Table 20.6
k Variables for samples

Sample	Age x_1	Gender x_2	TP x_3	Alb x_4	ChE x_5	LDH x_6	ALP x_7	GTP x_8	LAP x_9	Tch x_{10}	TG x_{11}	PL x_{12}	Cr x_{13}	BUN x_{14}	UA x_{15}
1	30	10	8	4	0.59	336	13.7	192	856	144	138	199	1.6	16	3.5
2	46	10	7	3.3	0.48	365	25.8	274	1027	113	176	218	1.9	18	4.5
3	22	1	7.2	3.8	0.52	255	10.6	123	704	141	190	206	1.4	11	3.5
4	31	1	7	3.8	0.57	723	24.3	152	841	125	45	164	1.3	15	3.1
5	33	10	6.7	3.4	0.51	281	14.4	96	613	89	133	147	1.4	11	1.9
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
34	71	1	6.8	3.3	0.54	533	22.7	156	646	130	126	220	1.6	13	2.6
35	61	1	7.2	4	0.53	747	13.7	123	971	174	91	215	1.4	13	4.4
36	16	1	6.7	3.9	0.51	458	12.8	37	495	122	151	201	1.9	11	3.6

Table 20.7
Normalized data

| Sample	Age z_1	Gender z_2	TP z_3	Alb z_4	ChE z_5	LDH z_6	ALP z_7	GTP z_8	LAP z_9	Tch z_{10}	TG z_{11}	PL z_{12}	Cr z_{13}	BUN z_{14}	UA z_{15}	MD Abnormal
1	−0.94	0.73	3.19	−2.09	−3.05	6.13	6.01	13.12	9.40	−1.80	1.53	0.34	2.30	0.31	−1.25	25.1
2	0.61	0.73	−0.19	−7.12	−4.20	7.39	14.93	19.47	12.46	−3.19	2.62	1.26	3.88	1.00	−0.43	66.5
3	−1.71	−1.36	0.49	−3.53	−3.78	2.60	3.73	7.78	6.67	−1.93	3.02	0.68	1.25	−1.43	−1.25	16.9
4	−0.84	−1.36	−0.19	−3.53	−3.26	23.01	13.82	10.02	9.13	−2.65	−1.14	−1.36	0.73	−0.04	−1.58	59.8
5	−0.65	0.73	−1.20	−6.40	−3.89	3.73	6.53	5.69	5.04	−4.28	1.38	−2.19	1.25	−1.43	−2.56	17.0
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
34	3.03	−1.36	−0.86	−7.12	−3.57	14.72	12.64	10.33	5.63	−2.43	1.18	1.36	2.30	−0.74	−1.99	41.8
35	2.06	−1.36	0.49	−2.09	−3.68	24.05	6.01	7.78	11.46	−0.45	0.18	1.12	1.25	−0.74	−0.52	52.8
36	−2.29	−1.36	−1.20	−2.81	−3.89	11.45	5.35	1.12	2.92	−2.79	1.90	0.43	3.88	−1.43	−1.17	25.0
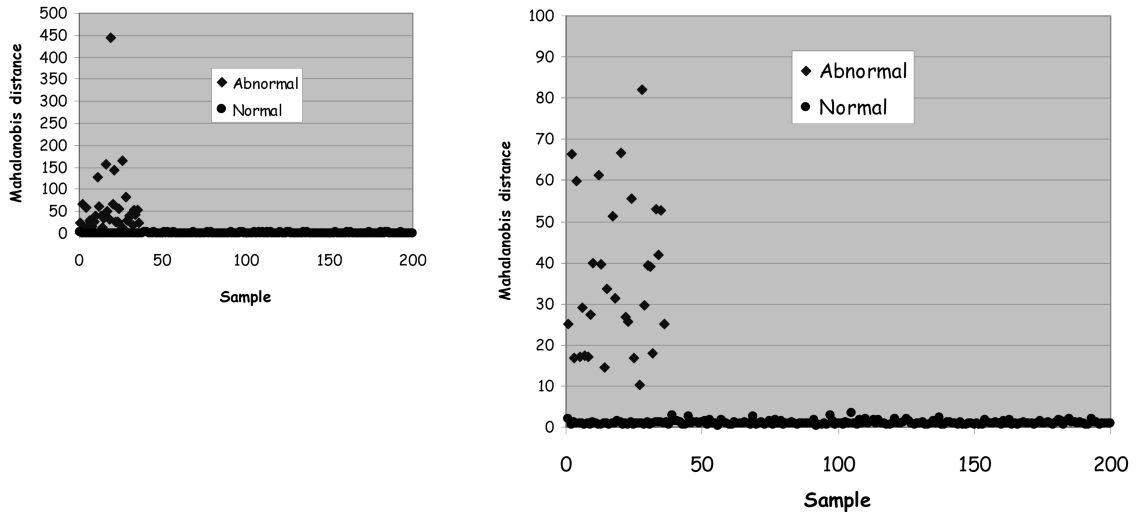
Figure 20.7
Discrimination power

Step 8. Optimize MTS.

Step 8.1. Assign variables to a two-level orthogonal array. Define level 1, to use variable, and level 2, not to use variable. Table 20.8 presents the two-level array. Recognize that run 1 of L_{16} is to use all 15 variables, and others are using different combinations of variables. The MTS is executed for all 16 runs, and data will be the MD for those 36 abnormal samples.

Step 8.2. Redo MTS for each run of the orthogonal array. First, calculate the MD for the abnormal sample as data and then use the SN ratio as an assessment criteria for discrimination power, as shown in Table 20.9. Recognize that the larger the MD values for abnormal samples, the better the discrimination is. In this example, number 4 shows very poor discrimination, and numbers 1, 7, and 16 show very good discrimination. The SN ratio for a larger-the-better response is used to assess the discrimination power. Ideally, a dynamic SN ratio where M is the true abnormality level and y is the computed MD value would be the ideal assessment.

Now we are ready to generate the response graph for the SN ratio (Figure 20.8).

Step 8.3. Make a response graph for the SN ratio. Evaluate how each variable contributes to discrimination power. Recognize that the higher the SN ratio, the better the discrimination power is. Notice that:

❏ Level 2 has a higher SN ratio for x_1, x_2, x_3, x_{14}, and x_{15}.

❏ Levels 1 and 2 have roughly the same performance for x_4, x_{11}, x_{12}, and x_{13}.

❏ Level 1 has higher SN ratios for x_5, x_6, x_7, x_8, x_9, and x_{10}.

From this we can conclude:

❏ x_1, x_2, x_3, x_{14}, and x_{15} are harming discrimination.

❏ x_4, x_{11}, x_{12}, and x_{13} do not contribute to discrimination.

❏ x_5, x_6, x_7, x_8, x_9, and x_{10} contributes greatly to discrimination.

Table 20.8
Two-level orthogonal array

L_{16}	x_1 1	x_2 2	x_3 3	x_4 4	x_5 5	x_6 6	x_7 7	x_8 8	x_9 9	x_{10} 10	x_{11} 11	x_{12} 12	x_{13} 13	x_{14} 14	x_{15} 15
1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1
2	1	1	1	1	1	1	1	2	2	2	2	2	2	2	2
3	1	1	1	2	2	2	2	1	1	1	1	2	2	2	2
4	1	1	1	2	2	2	2	2	2	2	2	1	1	1	1
5	1	2	2	1	1	2	2	1	1	2	2	1	1	2	2
6	1	2	2	1	1	2	2	2	2	1	1	2	2	1	1
7	1	2	2	2	2	1	1	1	1	2	2	2	2	1	1
8	1	2	2	2	2	1	1	2	2	1	1	1	1	2	2
9	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
10	2	1	2	1	2	1	2	2	1	2	1	2	1	2	1
11	2	1	2	2	1	2	1	1	2	1	2	2	1	2	1
12	2	1	2	2	1	2	1	2	1	2	1	1	2	1	2
13	2	2	1	1	2	2	1	1	2	2	1	1	2	2	1
14	2	2	1	1	2	2	1	2	1	1	2	2	1	1	2
15	2	2	1	2	1	1	2	1	2	2	1	2	1	1	2
16	2	2	1	2	1	1	2	2	1	1	2	1	2	2	1

L_{16}	x_1 1	x_2 2	x_3 3	x_4 4	x_5 5	x_6 6	x_7 7	x_8 8	x_9 9	x_{10} 10	x_{11} 11	x_{12} 12	x_{13} 13	x_{14} 14	x_{15} 15
1	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use	Use
2	Use	Use	Use	Use	Use	Use	Use								
3	Use	Use	Use					Use	Use	Use	Use				
4	Use	Use	Use									Use	Use	Use	Use
5	Use			Use	Use			Use	Use			Use	Use		
6	Use			Use	Use					Use	Use			Use	Use
7	Use					Use	Use	Use	Use					Use	Use
8	Use					Use	Use			Use	Use	Use	Use		
9		Use		Use		Use		Use		Use		Use		Use	
10		Use		Use		Use			Use		Use		Use		Use
11		Use			Use		Use	Use		Use			Use		Use
12		Use			Use		Use		Use		Use	Use		Use	
13			Use	Use			Use	Use			Use	Use			Use
14			Use	Use			Use		Use	Use			Use	Use	
15			Use		Use	Use		Use			Use		Use	Use	
16			Use		Use	Use			Use	Use		Use			Use

Table 20.9
Response data

	1	2	3	4	5	6	7	8	9	10	.	34	35	36	SN Ratio
1	25.1	66.5	16.9	59.8	17.0	29.2	17.3	17.2	27.3	39.9	.	41.8	52.8	25.0	28.39
2	17.3	54.4	10.2	112.9	21.5	12.2	13.1	21.4	26.1	33.0	.	60.2	95.6	34.6	23.40
3	34.4	77.1	18.6	30.4	12.1	38.0	21.1	10.1	22.1	51.1	.	27.0	29.4	8.5	21.93
4	3.9	4.2	2.8	1.9	3.6	2.6	5.3	2.8	5.4	2.2	.	5.4	2.1	9.2	7.90
5	38.7	90.8	23.2	32.2	22.4	42.7	25.8	24.8	31.4	63.6	.	34.0	34.7	11.7	25.45
6	4.1	15.8	9.5	4.4	13.4	8.5	7.9	14.9	10.3	7.4	.	10.6	3.0	8.2	15.57
7	39.8	93.4	16.3	106.5	16.0	34.7	23.2	14.4	32.3	59.1	.	59.2	93.8	26.9	28.44
8	16.6	63.9	13.4	107.7	16.7	21.9	9.0	14.8	29.2	29.2	.	62.2	88.8	37.4	26.47
9	38.0	100.7	22.2	96.1	23.2	34.5	18.8	24.3	38.0	35.1	.	67.1	92.4	32.2	27.35
10	21.0	42.1	12.7	88.2	15.9	35.2	21.9	17.5	17.3	53.6	.	43.4	95.7	24.3	27.11
11	43.5	108.5	21.3	50.9	20.9	32.2	22.8	19.0	36.9	50.2	.	50.8	21.2	16.4	24.18
12	24.1	61.5	17.5	47.6	15.3	40.4	24.3	14.2	18.0	68.1	.	30.4	37.2	9.9	25.04
13	40.3	100.0	15.5	49.0	20.6	25.6	23.1	22.3	24.7	51.2	.	45.9	16.4	5.2	18.47
14	23.4	68.4	13.8	43.9	22.6	40.2	25.3	24.6	15.1	62.3	.	37.8	27.5	11.5	24.07
15	37.8	79.2	16.1	95.0	12.6	22.6	14.9	10.8	35.1	35.4	.	51.5	96.3	30.0	25.33
16	32.2	66.8	23.4	96.8	19.7	54.6	26.1	17.1	35.7	58.3	.	54.0	107.9	34.6	30.70

Figure 20.8
SN ratios
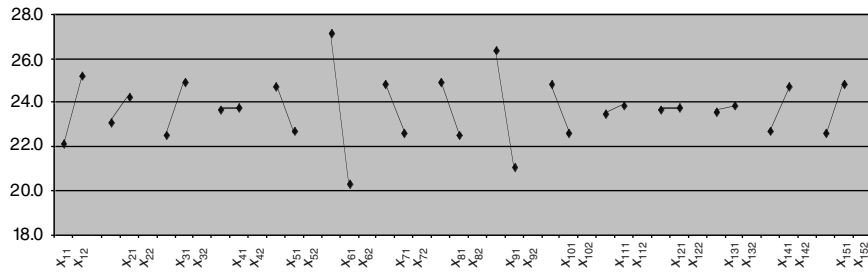
As of now our conclusion is as follows:

Optimization: Keep x_5, x_6, x_7, x_8, x_9, and x_{10}.

Discard all others.

As with all other Taguchi methods, we need to confirm this result.

Step 8.4. Confirm the optimum design. Now we evaluate the discrimination power under the optimum scheme. The result of confirmation using only x_5, x_6, x_7, x_8, x_9, and x_{10} is shown in Figure 20.9.

Conclusions: (1) discrimination has improved and (2) it only requires six variables. This new system has led to an improvement, as shown in Table 20.10. As can
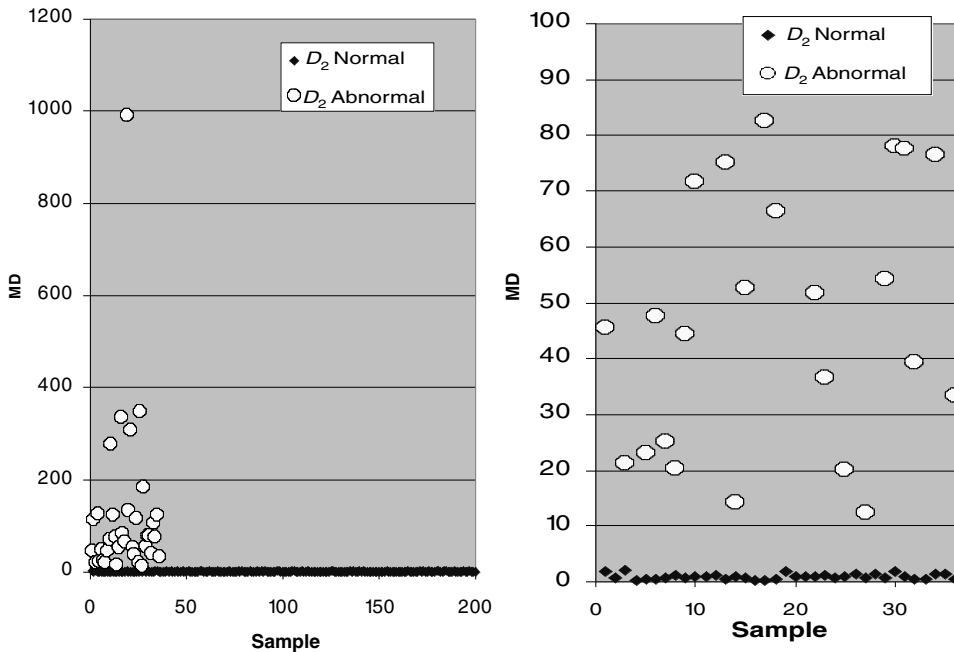
Figure 20.9
Result of confirmation

Table 20.10

Improvement in diagnosis of liver disease

	Traditional Method			Mahalanobis–Taguchi		
	Negative	**Positive**	**Total**	**Negative**	**Positive**	**Total**
Negative	28	**51**	79	63	**16**	79
Positive	**1**	15	16	**1**	15	16

be seen in Table 20.10, the current diagnosis method and Mahalanobis–Taguchi system have the same error rate for false negatives. The current diagnosis method has 65% false positives, requiring further testing, whereas the MTS has only a 20% false positive rate.

20.5. Mahalanobis–Taguchi Gram–Schmidt Process

There is also an alternative way of evaluating MD with the help of the Gram–Schmidt orthogonalization process (GSP). This process is described in the following section.

20.6. Gram–Schmidt's Orthogonalization Process

Given linearly independent vectors Z_1, Z_2, ... , Z_k, there exist mutually perpendicular vectors U_1, U_2, ... , U_k with the same linear span.

The Gram–Schmidt vectors are constructed sequentially by setting (Figure 20.10)

$$U_1 = Z_1$$

$$U_2 = Z_2 - \frac{Z_2' U_1}{U_1' U_1} U_1$$

$$\vdots$$

$$U_k = Z_k - \frac{Z_k' U_1}{U_1' U_1} U_1 - \cdots - \frac{Z_k' U_{k-1}}{U_{k-1}' U_{k-1}} U_{k-1}$$

where the prime denotes the transpose of a vector.

While calculating the MD using GSP, standardized values of the variables are used. Therefore, in the set of equations above, Z_1, Z_2, ... , Z_k correspond to

$Z_1, Z_2, . ., Z_k$ (original vectors) → GSP → $U_1, U_2, . ., U_k$ (orthogonal vectors)
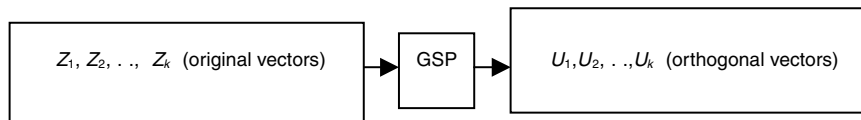
Figure 20.10
Gram–Schmidt's process

standardized values. From this set of equations it is clear that the transformation process depends largely on the first variable.

20.7. Calculation of MD Using the Gram–Schmidt Process

Let us suppose that we have a sample of size n and that each sample contains observations on k variables. After standardizing the variables, we will have a set of standardized vectors. Let these vectors be

$$Z_1 = (z_{11}, z_{12}, \ldots , z_{1n})$$

$$Z_2 = (z_{21}, z_{22}, \ldots , z_{2n})$$

$$\vdots$$

$$Z_k = (z_{k1}, z_{k2}, \ldots , z_{kn})$$

After performing GSP, we have orthogonal vectors as follows:

$$U_1 = (u_{11}, u_{12}, \ldots , u_{1n})$$

$$U_2 = (u_{21}, u_{22}, \ldots , u_{2n})$$

$$\vdots$$

$$U_k = (u_{k1}, u_{k2}, \ldots , u_{kn})$$

It can easily be followed that the mean of vectors $U_1, U_2, \ldots , U_k$ is zero. Let $s_1, s_2, \ldots , s_k$ be standard deviations of $U_1, U_2, \ldots , U_k$, respectively. Since we have a sample of size n, there will be n different MDs. MD corresponding to the jth observation of the sample is computed using the equation

$$MD_j = \frac{1}{k} \left(\frac{u_{1j}^2}{s_1^2} + \frac{u_{2j}^2}{s_2^2} + \cdots + \frac{u_{kj}^2}{s_k^2} \right)$$

where $j = 1, \ldots , n$. The values of MD obtained from the inverted matrix method are approximately equal to each other. Table 20.11 below shows MD values by the inverted matrix method and the MTGS method from the preceding example.

Table 20.11
MD values using MTGS and MTS

(a) Healthy Group		
	MD Values (1,2, ... , 200)	**Average MD**
MTGS	2.105694, 0.605969, ... , 0.899975, 0.765149, 0.913744	0.99499991
MTS	2.105699, 0.605892, ... , 0.899972, 0.765146, 0.913695	0.99498413
(b) Abnormal Group		
	MD Values (1,2, ... , 36)	**Average MD**
MTGS	25.13597, 66.46272, ... , 41.84671, 52.76765, 24.98069	60.09035
MTS	25.13559, 66.4609, ... , 41.84529, 52.76777, 24.98046	60.08957

20.8. MTS/MTGS Method versus Other Methods

MTS/MTGS methods are intended to provide meaningful information based on patterns. These patterns are identified based on several characteristics. The methods are different from classical multivariate methods in the sense that the methods used in MTS/MTGS are data analytic (i.e., they do not require any assumptions on the distribution of input variables) rather than being on probability-based inference. Artificial neural networks (ANN) also use data analytic methods for pattern recognition problems.

MTS/MTGS methods are different from classical multivariate methods in the following ways:

1. In these methods the Mahalanobis distance is suitably scaled and used as a measure of severity of various conditions.

2. The conditions outside the normal space (i.e., abnormal conditions) are unique and do not constitute a separate population.

The objectives of these multivariate/pattern recognition techniques are implicit in MTS/MTGS methods. These are in addition to the primary objective of developing a multidimensional measurement scale. Like principal component analysis (PCA), MTS/MTGS methods can be used for dimensionality reduction. In fact, dimensionality reduction is done in terms of original variables. In PCA, although dimensionality reduction is done by calculating principal components, to calculate one principal component we need all of the original variables.

Like the discrimination and classification method and regression analysis, MTS/MTGS methods can be used to find the "normals" and "abnormals." Further, they can be used to measure the degree of abnormality. The degree of abnormality cannot be measured by using the discrimination and classification method because its objective is to classify observations into different groups. It is possible to measure the degree of abnormality with regression analysis, but the method of doing this is complex.

Like stepwise regression or a test for additional information, MTS/MTGS methods can be used to screen the variables. However, stepwise regression is a probabilistic approach. The method does not guarantee the best subset regression model. As it terminates with one final equation, an inexperienced analyst may conclude that he or she has found the optimal subset of the model.

Like multivariate charts, MTS/MTGS methods can be used to monitor and control various process conditions based on their severity. It is to be noted that in process control charts the degree of abnormality is judged with respect to probabilistic control limits, whereas in MTS/MTGS the degree of abnormality is judged with respect to a threshold obtained from QLF.

20.9. MTS/MTGS versus Artificial Neural Networks

As compared with artificial neural networks (ANNs), MTS/MTGS methods are similar in the sense that they are data analytic. Using ANNs, it is not as easy to achieve dimensionality reduction as in the case of MTS/MTGS methods. With ANNs, if a new pattern is added in the system, the weight value of the network will be changed because the ANN has to recognize this pattern along with old patterns. Also, randomization plays an important role because ANNs have a

tendency to forget old patterns. So patterns are randomized with ANNs. In MTS/ MTGS methods, if a new pattern is added, the MS corresponding to this pattern has to be constructed and added to the system, and there is no need to randomize the patterns. ANNs will not help in finding the relationship between input and output. In MTS/MTGS methods, such relationships can be obtained. Another disadvantage of ANNs is setting the number of hidden layers, as this number would affect ultimate results. There is no definite means to set the number of hidden layers.

20.10. MTS Applications

MTS can be used for two major objectives: diagnosis (pattern recognition) and forecasting. Following are some of the potential areas of application.

Medical Diagnosis

In a conventional checkup, diagnosis is made by the experience of a physician along with the results of testing. There are usually many test items, and each item has a range. A person will be judged, and if the tested result of one item falls beyond the range, he or she will be sent to a secondary examiner, and a closer examination will be made by increasing the number of check items.

In biochemical testing, normal values, in general, are determined arbitrarily by the tester and test chemical manufacturers, and in extreme cases, textbook values are used without modification.

Many of the test items have various levels of correlations. In the application of MTS, test results from a healthy group are collected to construct a Mahalanobis space (normal space or reference space) for diagnosis. This approach can be used for all types of diagnosis. It can also be used to predict the time it will take a patient to recover.

Perhaps the most exciting potential of application is in research on medicines and medical treatments. Currently, a newly developed medicine is evaluated using a double-blind test with two groups. One group is given the test drug, and the other, a placebo. Not only does such a test require a large number of people, it is inhumane for the group taking the placebo. By the use of MTS, the change in Mahalanobis distance can detect the effect of the new drug or the new treatment method, possibly using only one person in a short period of time.

Manufacturing

In manufacturing, pattern recognition is widely used. For example, many inspections are done by eyeball observation, such as the appearance of welded or soldered component parts. When the fraction defective in a production line is very low, it is difficult for a worker to inspect 300 pieces an hour, and therefore easy to overlook a defective. Moreover, meeting specifications one by one does not guarantee that it is a good product because of correlations. Many successful studies of MTS have been reported from the mechanical, electrical, and chemical industries.

MTS is also very useful in monitoring a manufacturing process. For example, a semiconductor process can be very complex and consists of many variables, such as temperature, flow, and pressure. In that case, the normal group would be defined as the group existing when the process is producing perfect products.

In public buildings or hotels, fire alarm systems are installed by law. Perhaps most of us have experienced a false alarm: An alarm sounds but there is no fire, or a smoke detector sounds as a result of cigarette smoking or barbecuing. In the development of fire alarm systems, data such as temperature or the amount of smoke are collected from the situations without fire. These data are considered as a reference group for the construction of a Mahalanobis space.

Fire Detection

In the case of earthquakes, it is difficult to know what type of data should be collected. The method used for data collection depends on what type of forecasting is to be done. For example, we may try to forecast an earthquake one hour from now. But such forecasting may not be very useful except for being able to take such actions as shutting off the gas valve.

Earthquake Forecasting

Suppose that we are going to forecast an earthquake that will occur 24 to 48 hours from now. It is an interesting problem, and we must know in MTS that the data to be collected are under "normal" conditions. In medical checkups, the data are collected from the "healthy people." In fire detection, data are collected under "no fire" conditions. Similarly, the data "without earthquake" must be collected as a reference group. Weather bureaus have seismograph records for years. The amplitude of 48 hours without an earthquake is collected to construct a Mahalanobis space. Then the amplitudes during the 48 hours following an earthquake are collected and compared for the study of future forecasting.

MTS could be used for weather forecasting—not to discuss current forecasting methodology itself but to provide an approach that summarizes all types of existing data and to simplify the process by reducing information that is not contributing to a proper forecast.

Weather Forecasting

In designing air bags in a car, we want to avoid both types of error: an air bag actuated when there was no crash and an air bag that did not actuate at the moment of a crash. To construct a Mahalanobis space, acceleration at several locations in a car is recorded from time to time at an interval of a few thousandth of a second. Such data are collected at various driving conditions: driving on a very rough road, quick acceleration or quick stop, and so on. Those data are used with other sources of information, such as driving speed, to construct a Mahalanobis space. The computer in the air bag system is constantly calculating a Mahalanobis distance. When the distance becomes greater than a predicted threshold, the system sends a signal to actuate the air bag(s) before a crash. On the passenger side, you don't want to deploy the air bag when a child is sitting there or when an adult is sitting in the seat and bending over. Such MTS studies are currently being conducted within automotive companies.

Automotive Air Bag Deployment

Automotive manufacturers are developing sensors that detect a dangerous condition during driving. From various sensor outputs, we want to detect a situation immediately before an accident takes place so that the system can notify the driver or even take over the control of the car by braking, accelerating, and/or steering. Research is in the early stages.

Automotive Accident Avoidance Systems

Prediction is one of the most important areas in pattern recognition. In the business area, banks or credit card companies make predictions for loan or credit card

Business Applications

approval based on the application form or the information from credit companies. A Mahalanobis space can be formed from the existing good customers and the Mahalanobis distance used to predict creditworthy applicants.

Other Applications   MTS can be applied in many other areas, such as fingerprints, handwriting, or character or voice recognition. Although some products of this type are already in the market, they are far from mistake-free. The use of Mahalanobis distance could contribute to further improvement.

20.11. Some Important Considerations

The MTS basically provides a measurement system to assess abnormality. To define a measurement system, we need two parameters defined: the origin and a unit scale. For an MTS, the origin, zero, and one unit are defined by the normal group. The ideal function in measurement is y, the measured value, being proportional to M, the true value. Therefore, it is ideal to use a dynamic SN ratio where M is the true abnormality and y is the MD (or y is the square root the MD). The reciprocal of the SN ratio is the error variance of measurement. The procedure is to evaluate each sample in the abnormal group and assign a number that reflects the abnormality level. That number is used as its signal level. This evaluation and scoring must be done by a professional in the field (Figure 20.11.)

In a medical treatment center, for example, the Mahalanobis distance of a certain patient was calculated from medical test results before the patient was treated. After a certain period of treatment, another Mahalanobis distance was calculated from the new test results. If it is known that the larger the Mahalanobis distance, the more severe the illness, the value calculated from the new test should become smaller if the illness has improved. This is illustrated in Figure 20.12. This is a very important area in which MTS can be used for forecasting. From the relationship between the Mahalanobis distance and the time of treatment, a physician is able to predict when a patient will be able to leave the hospital.

In medical research, researching the effect of a new treatment or a new drug can be made using just one patient instead of relying on double-blind tests, where
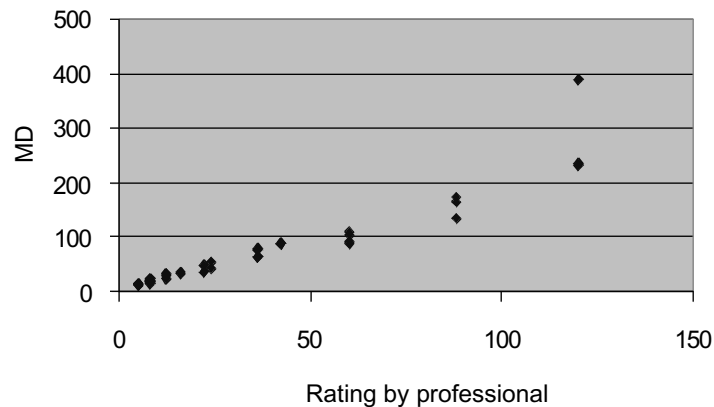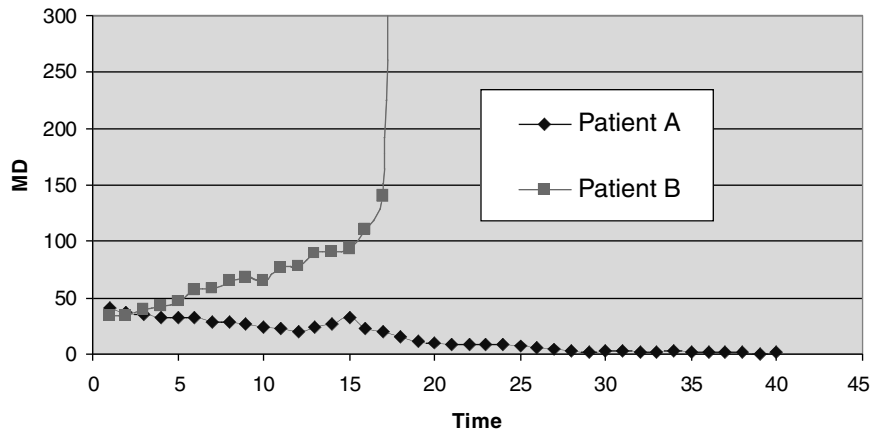
Figure 20.11
Rating of abnormality level

Figure 20.12
Monitoring abnormal condition

thousands of people are studied; sometimes such a study takes years. By use of the dynamic SN ratio, the trend can be observed in a short period of time. Such applications have a huge potential for future research and would bring immeasurable benefits to society.

Of course, it is all right to discuss two types of errors, or chi-square distribution, as a reference. But it is essential to realize that the objective of the MTS approach is to use the SN ratio for estimation of the error of misjudgment and also for forecasting.