Chapter 18
Quantitative Trait Associated Microarray Data Analysis

24

Quantitative Trait Associated Microarray Data Analysis

Differential expression analysis often applies to discrete phenotypes (primarily dichotomous phenotypes). The phenotype is often defined as"normal" or "affected". If a phenotype is measured quantitatively, it is often converted into two or a few discrete phenotype groups so that a differential expression analysis or an ANOVA method for multiple comparisons can be applied. It is obvious that such discretization will result in loss of information. The current microarray data analysis technique has not been able to efficiently analyze the association of gene expression with a continuous phenotype. Pearson correlation between gene expression and a continuous phenotype has been proposed.Blalock et al. (2004) ranked genes according to the correlation coefficients of gene expression with MMSE, a quantitative measurement of the severity of Alzheimer disease, and detected many genes that are associated with Alzheimer disease. Pearson correlation is intuitive and easy to calculate. However, it may not be optimal because (1) the correlation coefficient may not be the best indicator of the association, (2) higher order association cannot be detected, (3) data are analyzed individually with one gene at a time, and (4) the method cannot be extended to association study of gene expression with multiple continuous phenotypes. Potokina et al. (2004) investigated the association of gene expression with six malting quality phenotypes (quantitative traits) of ten barley cultivars. They compared the distance matrix of each gene expression among the ten cultivars with each of the distance matrix calculated from the phenotypes using the G-test statistic. The distance matrix comparison approach may have the same flaws as the correlation analysis. Recently, we proposed to use the regression coefficient of the expression on a continuous phenotype as the indicator of the strength of association (Jia and Xu, 2005). Instead of analyzing one gene at a time, we took a model-based clustering approach to studying all genes simultaneously. Qu and Xu (2006) extended the model-based clustering algorithm to captures genes with higher order association with the phenotype.

24.1 Linear association

24.1.1 Linear model

Let $Z = \begin{bmatrix} Z_1 \cdots Z_M \end{bmatrix}^T$ be the phenotypic values of a quantitative trait (a continuous variable) for M individuals who are also microarrayed for N genes. Let $y_j = \begin{bmatrix} y_{1j} \cdots y_{Mj} \end{bmatrix}^T$ be the expressions of the jth genes on all the M individuals for $j = 1, ..., N$ where N is the total number of genes. The linear model for gene expression associated with the phenotype is

$$y_j = 1\beta_j + Z\gamma_j + \varepsilon_j \tag{24.1}$$

where 1 is a vector of unity with dimension $M \times 1$, β_j is the intercept and γ_j is the regression coefficient representing the association of gene j with the phenotype under investigation. The residual error ε_j is an $M \times 1$ vector with an assumed $N(0, I\sigma^2)$ distribution. Since β_j is irrelevant to the association study, it can be eliminated from the model. The simplest way to eliminate β_j is to centralize the expression by $y_j = y_j - \bar{y}_j$, where

$$\bar{y}_j = \frac{1}{M}\sum_{k=1}^{M} y_{kj} \tag{24.2}$$

The phenotypic value should also be centralized using $Z = Z - \bar{Z}$ where

$$\bar{Z} = \frac{1}{M}\sum_{k=1}^{M} Z_k \tag{24.3}$$

The linear model for the centralized gene expression becomes

$$y_j = Z\gamma_j + \varepsilon_j \tag{24.4}$$

Through centralization, we have eliminated β_j from the model. We can now focus on γ_j because it represents the strength of the association of y_j with Z.

24.1.2 Cluster analysis

Clustering genes based on the regression coefficients of gene expressions on a quantitative trait was first proposed by Jia and Xu (2005). We now use a Gaussian mixture with C components to describe the regression coefficients,

$$p(\gamma_j) = \sum_{\kappa=1}^{C} \pi_\kappa N(\gamma_j|\mu_\kappa, \Sigma) \tag{24.5}$$

where μ_κ is the mean of cluster κ for $\kappa = 1, \cdots, C$ and Σ is a common variance (a single value, not a matrix). Again, the same EM algorithm described in the

time-course microarray data analysis can be applied here, except that $\psi(\tau)$ in the time-course microarray is now replaced by Z in the quantitative trait associated microarray data analysis.

The number of clusters C can be inferred using the BIC analysis (Schwarz, 1978). Most genes will be classified into the "neutral cluster" in which the cluster mean is close to zero.

24.1.3 Three-cluster analysis

We now use a Gaussian mixture with $C = 3$ components to describe the regression coefficients,

$$p(\gamma_j) = \sum_{\kappa=1}^{3} \pi_\kappa N(\gamma_j | \mu_\kappa, \Sigma) \tag{24.6}$$

where μ_κ is the mean of cluster κ for $\kappa = 1, 2, 3$ and Σ is a common variance. The means of the three clusters are restricted with $\mu_1 > 0$, $\mu_2 = 0$ and $\mu_3 < 0$. Under these restrictions, the neutral cluster is cluster 2 because $\mu_2 = 0$. All genes classified into this neutral cluster are neutral genes (not associated with the trait) while all other genes are differentially expressed or associated with the trait.

The usual EM algorithm is incapable of dealing with such a cluster analysis with constrained cluster means. Therefore, we adopted the stochastic EM algorithm. The SEM steps are similar to those described in the time-course microarray data analysis with the step of updating the cluster means modified to constrain the means within their defined domains. Let

$$\delta_j = \left[\delta_{j1} \ \delta_{j2} \ \delta_{j3} \right] \tag{24.7}$$

be the cluster indicator variables for the jth gene and

$$\rho_j = \left[\rho_{j1} \ \rho_{j2} \ \rho_{j3} \right] \tag{24.8}$$

be the corresponding posterior probabilities of δ_j . Note that the proportions of genes belonging to cluster κ is denoted by π_κ for all $\kappa = 1, 2, 3$. Let

$$\text{var}(y_j) = V = Z\Sigma^{-1}Z^T + I\sigma^2 \tag{24.9}$$

be the variance-covariance matrix of y_j . The cluster label δ_j is missing and thus it is sampled from its posterior distribution, a multinomial distribution of sample size one and a probability vector ρ_j. If we ignore the constraints, the posterior mean and posterior variance of μ_κ would be

$$\xi_\kappa = \left(\pi_\kappa N Z^T V^{-1} Z \right)^{-1} \sum_{j=1}^{N} \delta_{j\kappa} Z^T V^{-1} y_j \tag{24.10}$$

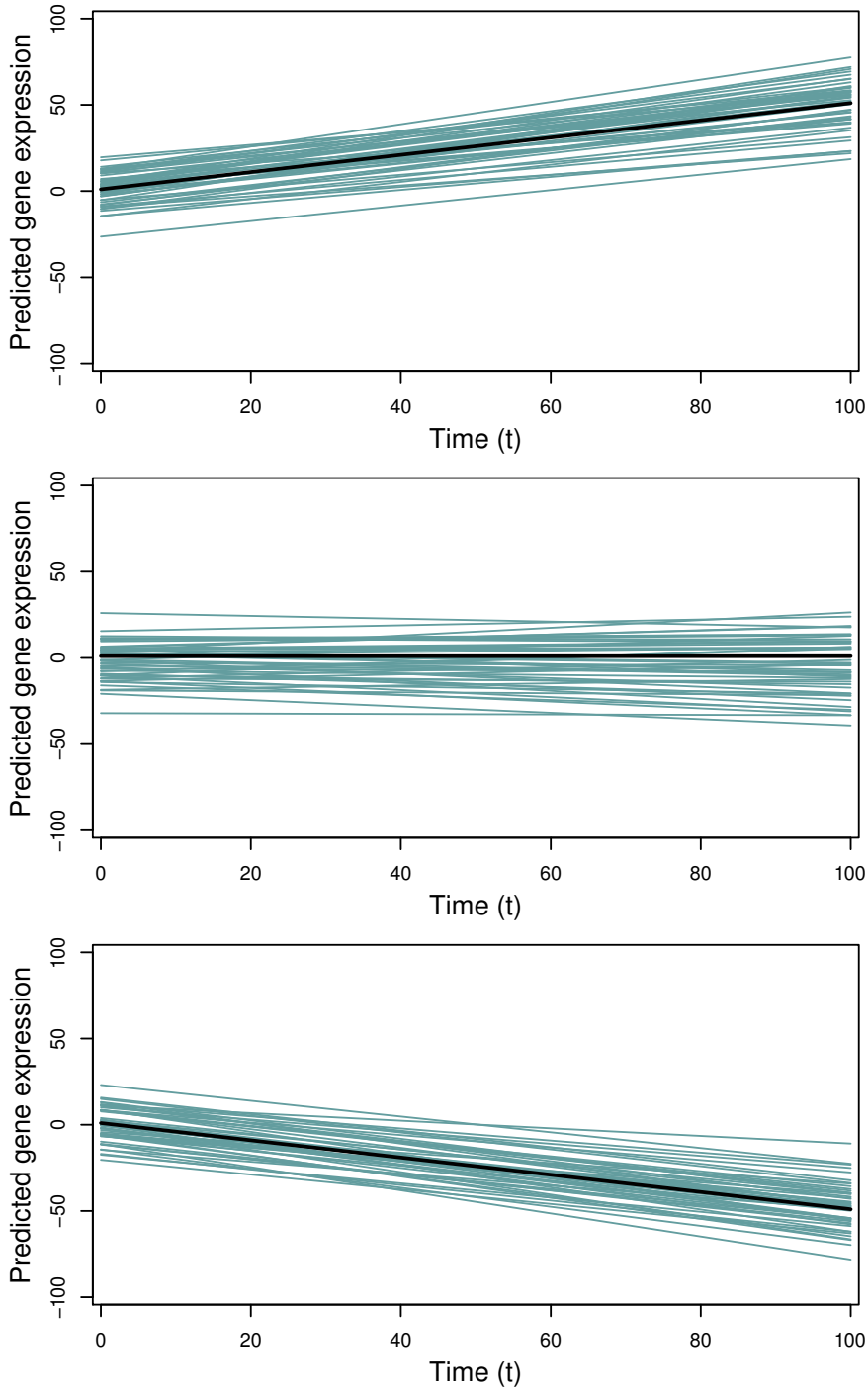

Fig. 24.1. Predicted expressions of 50 genes from each of the three designated clusters with $\mu_1 = 0.5$, $\mu_2 = 0$, $\mu_3 = -0.5$ and $\Sigma = 0.01$. The lines in green color represent the predicted expressions for individual genes and the line in black within each cluster represents the mean profile of the cluster.

and

$$\varphi_\kappa^2 = \left(\pi_\kappa N Z^T V^{-1} Z\right)^{-1} \tag{24.11}$$

respectively, for $\kappa = 1, 3$. The constrained estimate of $\mu_1 > 0$ is

$$\mu_1 = \xi_1 + \varphi_1 \frac{\phi\left(\frac{0-\xi_1}{\varphi_1}\right) - \phi\left(\frac{\infty-\xi_1}{\varphi_1}\right)}{\Phi\left(\frac{\infty-\xi_1}{\varphi_1}\right) - \Phi\left(\frac{0-\xi_1}{\varphi_1}\right)} = \xi_1 + \varphi_1 \frac{\phi\left(\frac{\xi_1}{\varphi_1}\right)}{\Phi\left(\frac{\xi_1}{\varphi_1}\right)} \tag{24.12}$$

where $\phi(x)$ and $\Phi(x)$ are the standardized normal density and the standardized normal function, respectively. The constrained estimate for $\mu_3 < 0$ is obtained through

$$\mu_3 = \xi_3 + \varphi_3 \frac{\phi\left(\frac{-\infty-\xi_3}{\varphi_3}\right) - \phi\left(\frac{0-\xi_3}{\varphi_3}\right)}{\Phi\left(\frac{0-\xi_3}{\varphi_3}\right) - \Phi\left(\frac{-\infty-\xi_3}{\varphi_3}\right)} = \xi_3 - \varphi_3 \frac{\phi\left(\frac{\xi_3}{\varphi_3}\right)}{1 - \Phi\left(\frac{\xi_3}{\varphi_3}\right)} \tag{24.13}$$

Equations (24.12) and (24.13) were derived following the theory of truncated normal distribution given by Cohen (1991). By definition, the mean of the neutral cluster is always $\mu_2 = 0$ and no estimation is required for μ_2. Originally, the class label for the jth gene was denoted by η_j, which is assigned $\eta_j = \kappa$ if gene j belongs to cluster κ, for $\kappa = 1, 2, 3$. The η_j variable was eventually converted into a 1×3 vector δ_j, a multinomial variable with sample size one. Given $\eta_j = \kappa$, we now rewrite the linear model in the form of a mixed model

$$y_j|_{\eta_j=\kappa} = Z\mu_\kappa + Z\alpha_j + \varepsilon_j \tag{24.14}$$

and perform the SEM iterations by sampling δ_j and updating the parameters. The SEM steps are summarized as follows.

1. Initialize all parameters within their legal domains.
2. Calculate the posterior probability that gene j belonging to cluster κ using

$$\rho_{j\kappa} = \frac{\pi_\kappa N(y_j|Z\mu_\kappa, V)}{\sum_{\kappa'=1}^{3} \pi_{\kappa'} N(y_j|Z\mu_{\kappa'}, V)} \tag{24.15}$$

3. Sample δ_j from

$$p(\delta_j) = \text{Multinomial}(\delta_j|1, \rho_j) \tag{24.16}$$

4. Update the cluster means μ_κ using the means of the truncated normal distributions given in equation (24.12) for μ_1 and equation (24.13) for μ_3 while forcing $\mu_2 = 0$.
5. Calculate the posterior mean and posterior variance for α_j using

$$\hat{\alpha}_j = \mathtt{E}(\alpha_j|\eta_j = \kappa) = \Sigma Z^T V^{-1} \delta_{j\kappa}(y_j - Z\mu_\kappa) \tag{24.17}$$

and

$$\hat{S}_j = \text{var}(\alpha_j|\eta_j = \kappa) = \Sigma - \Sigma Z^T V^{-1} Z \Sigma \tag{24.18}$$

6. Update the common variance of all clusters using

$$\Sigma = \frac{1}{N}\sum_{j=1}^{N}\sum_{\kappa=1}^{3}\delta_{j\kappa}\left(\hat{\alpha}_j^2 + \hat{S}_j\right) \tag{24.19}$$

7. Update the residual error variance using

$$\sigma^2 = \frac{1}{MN}\sum_{j=1}^{N}\sum_{\kappa=1}^{3}\delta_{j\kappa}y_j^T(y_j - Z\mu_\kappa - Z\hat{\alpha}_j) \tag{24.20}$$

8. Update the proportion of genes for each cluster using

$$\pi_\kappa = \frac{1}{N}\sum_{j=1}^{N}\delta_{j\kappa} \tag{24.21}$$

9. Repeat Steps 2 to 8 until all parameters converge to their stationary distributions.

After the SEM analysis, genes will be classified based on their posterior distributions of the clusters, i.e., gene j will be classified into cluster κ if

$$\rho_{j\kappa} = \max(\rho_{j1}, \rho_{j2}, \rho_{j3}) \tag{24.22}$$

Genes classified into the neutral cluster, $\kappa = 2$, will be excluded from the list of associated genes. Assume that gene j is classified into cluster κ, the predicted expression for gene j is calculated via

$$\hat{y}_j|_{\eta_j=\kappa} = Z(\hat{\mu}_\kappa + \hat{\alpha}_j) \tag{24.23}$$

where $\hat{\mu}_\kappa$ and $\hat{\alpha}_j$ are the estimated values obtained from the SEM analysis. Figure 24.1 illustrates the predicted gene expressions for the three designated clusters with $\mu_1 = 0.5$, $\mu_2 = 0$, $\mu_3 = -0.5$ and $\Sigma = 0.01$.

24.1.4 Differential expression analysis

An alternative method to detect genes associated with quantitative trait is through the differential expression analysis. This time, we use a Gaussian mixture with two components to model the distribution of the regression coefficients,

$$p(\gamma_j) = \pi N(\gamma_j|0, \Sigma_1) + (1-\pi)N(\gamma_j|0, \Sigma_0) \tag{24.24}$$

where both clusters have the same mean (zero value) but the two clusters have different cluster variances, Σ_0 and Σ_1. Let cluster 0 be the neutral cluster and cluster 1 be the differentially expressed cluster where Σ_1 is treated as a parameter and Σ_0 is set to a small constant, say $\Sigma_0 = 10^{-5}$. The proportion of genes coming from cluster 1 is denoted by π. The distributions of the

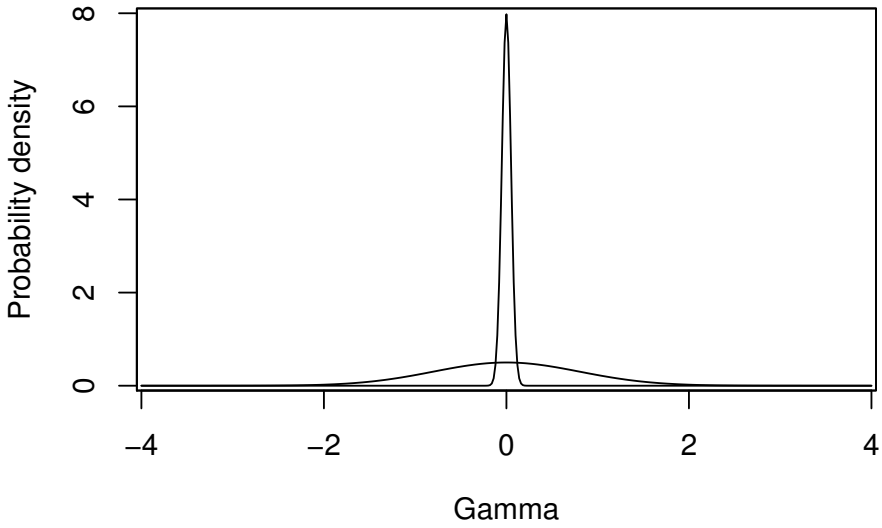

Fig. 24.2. Gaussian mixture with two components. The sharp curve represents $N(\gamma|0, \Sigma_0)$, the density of cluster 0, and the flat curve represents $N(\gamma|0, \Sigma_1)$, the density of cluster 1.

two components of the Gaussian mixture is illustrated in Figure 24.2. All genes with γ classified into the distribution represented by the flat curve are associated with the trait while the remaining genes (with γ classified into the distribution represented by the sharp curve) are neutral genes.

Under this model, the usually EM algorithm works well for parameter estimation. However, the SEM algorithm is preferred because of the high chance of finding the global maximum likelihood estimates of the parameters. Let η_j be an indicator variable with a value of one if gene j belongs to cluster 1 and zero if the gene belongs to cluster 0. Let ρ_j be the posterior probability of $\eta_j = 1$. Define

$$\text{var}(y_j|\eta_j) = V_j = Z\left[\eta_j\Sigma_1 + (1-\eta_j)\Sigma_0\right]Z^T + I\sigma^2 \tag{24.25}$$

as the variance-covariance matrix of y_j. The SEM steps are summarized as follows.

1. Initialize all parameters within their legal domains.
2. Calculate the posterior probability that gene j belonging to cluster 1 using

$$\rho_j = \frac{\pi N(y_j|0, \Gamma_1)}{\pi N(y_j|0, \Gamma_1) + (1-\pi)N(y_j|0, \Gamma_0)} \tag{24.26}$$

where

$$\Gamma_1 = Z\Sigma_1^{-1}Z^T + I\sigma^2 \tag{24.27}$$

and

$$\Gamma_0 = Z\Sigma_0^{-1}Z^T + I\sigma^2 \tag{24.28}$$

3. Sample η_j from the following Bernoulli distribution

$$p(\eta_j) = \text{Bernoulli}(\eta_j|\rho_j) \tag{24.29}$$

4. Calculate the posterior mean and posterior variance for γ_j using

$$\hat{\gamma}_j = \mathtt{E}(\gamma_j|\cdots) = \Theta_j Z^T V_j^{-1} y_j \tag{24.30}$$

and

$$\hat{S}_j = \text{var}(\gamma_j|\cdots) = \Theta_j - \Theta_j Z^T V_j^{-1} Z\Theta_j \tag{24.31}$$

where

$$\Theta_j = \eta_j\Sigma_1 + (1-\eta_j)\Sigma_0 \tag{24.32}$$

5. Update the variance of cluster one using

$$\Sigma_1 = \frac{1}{\pi N}\sum_{j=1}^{N}\eta_j \mathtt{E}(\gamma_j^2) = \frac{1}{\pi N}\sum_{j=1}^{N}\eta_j\left(\hat{\gamma}_j^2 + \hat{S}_j\right) \tag{24.33}$$

6. Update the residual error variance

$$\sigma^2 = \frac{1}{MN}\sum_{j=1}^{N} y_j^T(y_j - \eta_j Z\hat{\gamma}_j) \tag{24.34}$$

7. Update the proportion of genes for cluster 1 using

$$\pi = \frac{1}{N}\sum_{j=1}^{N}\eta_j \tag{24.35}$$

8. Repeat Steps 2 to 7 until all parameters converge to their stationary distributions.

After the SEM analysis, genes will be classified based on their posterior distributions of the clusters, i.e., gene j will be classified into cluster 1 if $\rho_j \geq 0.9$. All genes classified into cluster 1 will be declared as being associated with the phenotype.

Before we proceed to the next section, a comment on the second step of the SEM algorithm is helpful to inexperienced students. This step is used to calculate the posterior probability of gene j belonging to cluster 1 (the differentially expressed cluster). One would have thought to use the following equation

$$\rho_j = \frac{\pi N(\gamma_j|0,\Sigma_1)}{\pi N(\gamma_j|0,\Sigma_1) + (1-\pi)N(\gamma_j|0,\Sigma_0)} \tag{24.36}$$

because it is simpler than the one used in Step 2. However, the densities of the two distributions for γ_j requires value of γ_j, which is a missing quantity. We cannot simply replace the missing γ_j by the conditional expectation like we did for the other quantities that involve the missing γ_j. Therefore, we cannot use any probability densities containing missing parameters to calculate the conditional posterior probabilities of clustering assignment. In the next section, we will discuss the MCMC implemented Bayesian method, in which γ_j will be sampled. With the sampled γ_j, we can use the distributions of γ_j to calculate ρ_j.

24.2 Polynomial and B-spline

The linear association analysis cannot detect genes that are associated with the trait in higher orders. Although most associated genes may show linear relationship with the trait, some genes may show non-linear association with the trait. Figure 24.3 shows various forms of associations of genes with a quantitative trait. The polynomial and B-spline analysis can be used for detecting these genes. The procedure is identical to that described in the time-course microarray data analysis by replacing the time points with the phenotypic values of the quantitative trait, and thus no further discussion will be given here for the EM and SEM algorithm. The intercepts can be included in the analysis or excluded from the analysis via centralization. This section will focus on the Bayesian analysis implemented via the MCMC algorithm. In addition, we will deal with differential expression analysis, in which only two clusters are considered, one is the neutral cluster and the other is the differentially expressed cluster.

Let ζ be the scaled phenotypic value ranging from -1 to $+1$ and d be the degree of the polynomial to be fit. The linear model is

$$y_j = \psi(\zeta)\gamma_j + \varepsilon_j \tag{24.37}$$

where $\psi(\zeta)$ is an $M \times d$ orthogonal polynomial coefficient matrix and

$$\gamma_j = \left[\gamma_{1j} \cdots \gamma_{dj} \right]^T \tag{24.38}$$

are the regression coefficients. Again, we use a Gaussian mixture with two components to describe the regression coefficients,

$$p(\gamma_j) = \pi N(\gamma_j|0, \Sigma_1) + (1-\pi)N(\gamma_j|0, \Sigma_0) \tag{24.39}$$

where both clusters have a mean zero but with cluster specific variance-covariance matrices, Σ_1 and Σ_0, each is a $d \times d$ matrix. Let cluster 0 be the neutral cluster and cluster 1 be the differentially expressed cluster where Σ_1 is treated as a parameter matrix and $\Sigma_0 = 10^{-5} I_{d\times d}$ be a constant matrix with diagonal values taking a value of virtually zero. Again, let η_j be the

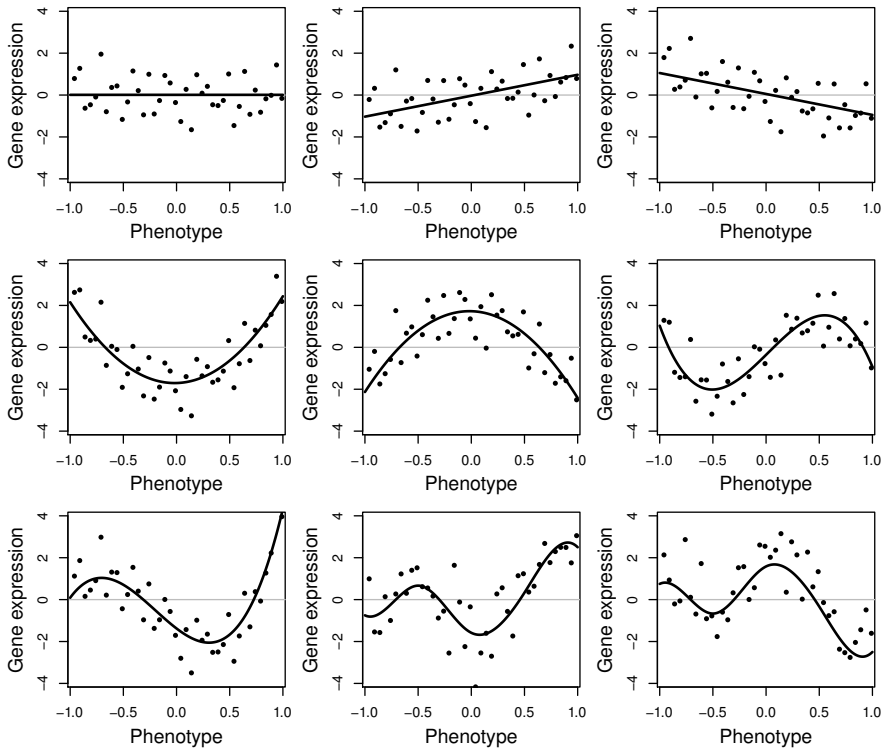

Fig. 24.3. Various forms of associations of gene expressions with a quantitative trait.

cluster label and ρ_j be the posterior probability of gene j coming from cluster 1. Define

$$\Theta_j = \eta_j \Sigma_1 + (1 - \eta_j)\Sigma_0 \tag{24.40}$$

and

$$\text{var}(y_j|\eta_j) = V_j = \psi(\zeta)\Theta_j\psi(\zeta)^T + I\sigma^2 \tag{24.41}$$

Under the Bayesian framework, we need to assign prior distributions to Σ_1 and σ^2. The residual error variance is assigned a scaled inverse chi-square distribution,

$$p(\sigma^2) = \text{Inv} - \chi^2(\sigma^2|\tau, \omega) \tag{24.42}$$

where $\omega = \tau = 0$ (the Jeffreys' prior). The covariance matrix Σ_1 is assigned a multivariate version of the scaled inverse chi-square distribution named the inverse Wishart distribution,

$$p(\Sigma_1) = \text{Inv} - \text{Wishart}(\Sigma_1|\tau, \omega) \tag{24.43}$$

where $\tau > d - 1$ is a prior degree of belief and $\omega > 0$ is a $d \times d$ scale matrix. We simply set $\tau = d$ and $\omega = 10^{-5} I_{d\times d}$. Another parameter in the analysis is

the mixing proportion denoted by π. A Beta prior is assigned to π,

$$p(\pi) = \text{Beta}(\pi|1,1) \tag{24.44}$$

These priors are conjugate and thus the posterior distributions of these parameters have the same forms of distributions as the priors.

The MCMC sampling process is summarized as follows.

1. Sample all unknown variables and parameters from their prior distributions.
2. Sample η_j from Bernoulli distribution,

$$p(\eta_j|\cdots) = \text{Bernoulli}(\eta_j|1,\rho_j) \tag{24.45}$$

   where ρ_j is a posterior probability calculated using

$$\rho_j = \frac{\pi N(\gamma_j|0,\Sigma_1)}{\pi N(\gamma_j|0,\Sigma_1) + (1-\pi)N(\gamma_j|0,\Sigma_0)} \tag{24.46}$$

3. Sample γ_j from its posterior distribution, which is multivariate normal

$$p(\gamma_j|\cdots) = N\left[\gamma_j|\Theta_j\psi(\zeta)^T V_j^{-1} y_j, \Theta_j - \Theta_j\psi(\zeta)^T V_j^{-1}\psi(\zeta)\Theta_j\right] \tag{24.47}$$

4. Sample the variance matrix from

$$p(\Sigma_1|\cdots) = \text{Inv} - \text{Wishart}\left(\Sigma_1 \,\middle|\, \tau + \sum_{j=1}^{N}\eta_j, \omega + \sum_{j=1}^{N}\eta_j\gamma_j\gamma_j^T\right) \tag{24.48}$$

5. Sample the residual error variance from

$$p(\sigma^2|\cdots) = \text{Inv}{-}\chi^2\left[\sigma^2 \,\middle|\, \tau + NM, \omega + \sum_{j=1}^{N} y_j^T(y_j - \eta_j\psi(\zeta)\gamma_j)\right] \tag{24.49}$$

6. Sample the mixture proportion from

$$p(\pi|\cdots) = \text{Beta}\left(\pi \,\middle|\, 1 + \sum_{j=1}^{N}\eta_j, 1 + N - \sum_{j=1}^{N}\eta_j\right) \tag{24.50}$$

7. Repeat Steps 2 to 6 until a desired length of the Markov chain is reached.

After the post-MCMC analysis (burn-in deletion and autocorrelation thinning), genes will be classified based on their posterior distributions of the clusters, i.e., gene j will be classified into cluster 1 if $\rho_j > 0.9$. All genes classified into cluster 1 will be declared as being associated with the phenotype.

Further analysis may be conducted on the Bayesian estimate of γ_j for all the differentially expressed genes. For example, we can use the K-means method to cluster all the γ_j into a few clusters. Different clusters represent different patterns (curves) of the expression profiles.

The MCMC algorithm developed for polynomial analysis also applies to B-spline analysis. The only difference is the dimension of the model. In the polynomial analysis, the dimension of γ_j is $d \times 1$ and the dimension for $\psi(\zeta)$ is $M \times d$. In the B-spline analysis, the dimension for γ_j is $p \times 1$ and the dimension for $\psi(\zeta)$ is $M \times p$, where $p = d + s + 1$ and d is the degree of the piecewise polynomial and s is the number of internal knots.

24.3 Multiple trait association

It is possible to study the association of genes with multiple quantitative traits. Let

$$Z_r = \left[Z_{r1} \cdots Z_{rM} \right]^T \tag{24.51}$$

be the phenotypic values of trait r collected from all the M individuals microarrayed. Denote the regression coefficient of gene j on the rth trait by γ_{rj} for $r = 1, \cdots, T$, where T is the number of traits. Assume that all gene expressions have been centralized and the phenotypic values of all the T traits have be standardized (centralized and normalized). The linear model for y_j can be expressed as

$$y_j = \sum_{r=1}^{T} Z_r \gamma_{jr} + \varepsilon_j = Z\gamma_j + \varepsilon_j \tag{24.52}$$

where

$$Z = \left[Z_1 \cdots Z_T \right] \tag{24.53}$$

is an $M \times T$ matrix and

$$\gamma_j = \left[\gamma_{j1} \cdots \gamma_{jT} \right]^T \tag{24.54}$$

be a $T \times 1$ vector.

Cluster analysis can be used to classify all genes into different clusters, depending on their associations with the multiple traits. Methods for the time-course microarray analysis and the single trait polynomial and/or B-spline association study apply to the multiple trait association study. The only difference is in the notation where $\psi(\zeta)$ in the polynomial analysis is replaced by matrix Z in the multiple trait analysis. Note that all the associations are linear in the multiple trait analysis. Extension to non-linear association of multiple traits is possible but it is difficult to implement. One has to define the degree of polynomial for each trait and different traits may be modeled using different degrees. If the degrees of polynomials are the same for all traits, the total number of effects in the γ_j vector is $T \times d$ for polynomial analysis and $T \times p$ for B-spline analysis where $p = d + s + 1$.