In Chapter 8 we discussed additive models (AM) of the form
\[ E(Y | X) = c + \sum_{\alpha=1}^{d} g_{\alpha}(x_{\alpha}) \] (9.1)

Note that we put \( EY = c \) and \( E\{g_{\alpha}(X_{\alpha})\} = 0 \) for identification.

In this chapter we discuss variants of the model (9.1). Recall that the main advantage of an additive model is that (compared to a fully nonparametric model) it avoids the curse of dimensionality and that the component functions are easy to interpret. Stone (1986) proved that this also holds for generalized additive models.

We will focus here on the modification of (9.1) by additional parametric, in particular linear, components. The resulting additive partial linear model (APLM) can be written as
\[ E(Y | U, T) = c + U^{\top} \beta + \sum_{\alpha=1}^{q} g_{\alpha}(T_{\alpha}) \] (9.2)

where we again use the partitioning \( X = (U, T) \) introduced in Chapter 7. Moreover, we will extend AM and APLM by a possibly nontrivial link function \( G \). This leads to the generalized additive model (GAM)
\[ E(Y | X) = G \left\{ c + \sum_{\alpha=1}^{d} g_{\alpha}(X_{\alpha}) \right\} \] (9.3)

or to the generalized additive partial linear model (GAPLM)
\[ E(Y | U, T) = G \left\{ c + U^{\top} \beta + \sum_{\alpha=1}^{q} g_{\alpha}(T_{\alpha}) \right\} \] (9.4)
It is obvious that model (9.4) is a modification of the GPLM. Consequently we will base the estimation of the GAPLM to a significant extent on the estimation algorithms for the GPLM. In fact, most of the methodology introduced in this chapter comes from the combination of previously introduced techniques for GLM, GPLM, and AM. We will therefore frequently refer to the previous Chapters 5, 7 and 8.

9.1 Additive Partial Linear Models

The APLM can be considered as a modification of the AM by a parametric (linear) part or as a nontrivial extension of the linear model by additive components. The main motivation for this partial linear approach is that explanatory variables are often of mixed discrete-continuous structure. Apart from that, sometimes (economic) theory may guide us how some effects inside the regression have to be modeled. As a positive side effect we will see that parametric terms can be typically estimated more efficiently.

Consider now the APLM from (9.2) in the following way:

\[
Y = c + \mathbf{U}^\top \beta + \sum_{a=1}^{q} g_a(T_a) + \varepsilon \quad (9.5)
\]

with \( E(\varepsilon|\mathbf{U}, T) = 0 \) and \( \text{Var}(Y|\mathbf{U}, T) = \sigma^2(\mathbf{U}, T) \). We have already presented estimators for \( \beta \) and \( m = c + \sum g_a \) in Section 7.2.2:

\[
\hat{\beta} = \left\{ \mathbf{U}^\top (I - S)^2 \mathbf{U} \right\}^{-1} \mathbf{U}^\top (I - S)^2 Y, \quad (9.6)
\]

\[
\hat{m} = S(Y - \bar{U}\hat{\beta}). \quad (9.7)
\]

As known from Denby (1986) and Speckman (1988), the parameter \( \beta \) can be estimated at \( \sqrt{n} \)-rate. The problematic part is thus the estimation of the additive part, as the estimate of \( m \) lacks precision due to the curse of dimensionality. When using backfitting for the additive components, the construction of a smoother matrix is principally possible, however, eluding from any asymptotic theory.

For this reason we consider a procedure based on marginal integration, suggested first by Fan, Härdle & Mammen (1998). Their idea is as follows: Assume \( \mathbf{U} \) has \( M \) realizations being \( \mathbf{u}^{(1)}, \ldots, \mathbf{u}^{(M)} \). We can then calculate the estimates \( \hat{g}_a^k \) for each of the subsamples \( k = 1, \ldots, M \). Now we average over the \( \hat{g}_a^k \) to obtain the final estimate \( \hat{g}_a \).

Note that this subsampling affects only the pre-estimation, when determining \( \hat{m} \) at the points over which we have to integrate. To estimate