into $T^\top DT$. Often it will be necessary to compute the pseudoinverse [131], since K^n tends to be rank degenerate for many practical kernels. Overall, the calculation costs $O((n+N)^2(n^*+N))$ operations.

In the case that x_l is an in-bound Support Vector we obtain a similar expression, the only difference being that the row and columns corresponding to x_l were removed from (12.87) in both K^n and Ψ^n. Recall that (see [337], 9.11.3.2a)

$$\begin{bmatrix} A & C \\ C^\top & D \end{bmatrix}^{-1} = \begin{bmatrix} A^{-1} + A^{-1}C(D - C^\top A^{-1}C)^{-1}C^\top A^{-1} & -A^{-1}C(D - C^\top A^{-1}C)^{-1} \\ -(D - C^\top A^{-1}C)^{-1}C^\top A^{-1} & (D - C^\top A^{-1}C)^{-1} \end{bmatrix} \tag{12.88}$$

By setting A equal to the square matrix in (12.87), with the contributions of x_l removed, and, further, identifying C with the remaining column vector (which contains the contribution of x_l) we see that (12.87) can be rewritten as in (12.83). ■

Note that we use $f(x) = \sum_i \alpha_i k(x_i, x)$.

Remark 12.19 (Modifications for Classification) *In the case of classification, the function expansions are usually given by sums of $y_i\alpha_i k(x_i, x)$. Simply replace α_i by $y_i\alpha_i$ throughout to apply Proposition 12.18.*

Since the assumptions regarding the stability of the types of SVs that led to this result are the same as the ones that led to (12.48) and (12.57), it comes as no surprise that the trick (12.88) can also be applied to compute (12.44) under those assumptions. This is due to the fact that in this case, the box constraints in 12.45 can be dropped. These issues are discussed in detail in [101].

In order to apply a similar reasoning to ν-SVM a slightly modified approach is needed. We only state the result for classification. The proof and extensions to regression and novelty detection are left as an exercise (see Problem 12.4).

Proposition 12.20 (Mean Field Leave-One-Out for ν-Classification) *Let K^n denote the $n \times n$ submatrix of kernel functions between in-bound SVs, and K^{n^*} the $(n^* - n) \times n$ submatrix of kernel functions between in-bound and bound SVs. Moreover, denote by $y^n \in \mathbb{R}^n$ the vector of labels (± 1) of the in-bound SVs, likewise by $y^{n^*} \in \mathbb{R}^{n^*-n}$ the vector of labels of bound SVs, and by $1^n \in \mathbb{R}^n$ and $1^{n^*} \in \mathbb{R}^{n^*-n}$ the corresponding vectors with all entries set to 1. Then, if x_l is an in-bound SV, the following approximation holds;*

$$(y_l f(x_l) - \rho) - (y_l f^l(x_l) - \rho^l) \approx \alpha_l y_l \left[\begin{bmatrix} K^n & y^n & 1^n \\ y^{n\top} & 0 & 0 \\ 1^{n\top} & 0 & 0 \end{bmatrix}^{-1}_{ll} \right]^{-1} \tag{12.89}$$