Rank-1 Updating Problem in MARS

April 16, 2024

We seek a representation in basis functions of the data $X \in \mathbb{R}^{n \times d}$, where n is the number of data points and d is the number of independent data dimensions, and $y \in \mathbb{R}^{n \times 1}$, which is the dependent data variable per data point.

Precisely, we seek the operator A, which minimises

$$\min_{A} \|A(X) - y\|_2^2 \tag{1}$$

, while maintaining some constraints on A, which are not relevant for this discussion but are maintainted by a superposition of basis functions.

To do so we explore different basis functions and have to solve a regression problem repeatedly. The problem lies in doing this efficiently. We have a data matrix $B \in \mathbb{R}^{n \times m}$, where m is the number of basis functions to explain the data and the regression problem is

$$\min_{x} \|Ba - y\|_2^2 \tag{2}$$

, where $a \in \mathbb{R}^{m \times 1}$ are the basis coefficients. The method of choice is solving the normal equations

$$B^T B a = B^T y \tag{3}$$

, where $B^T B$ are centered to have zero-mean, which results in the new problem

$$Va = c \tag{4}$$

with

$$V_{ij} = \sum_{k=1}^{n} B_{jk} \left[B_{ik} - \overline{B}_i \right] \tag{5}$$

$$c_i = \sum_{k=1}^{n} (y_k - \overline{y}) B_{ik} \tag{6}$$

, where $\overline{B}_i$ and $\overline{y}$ are the mean of the i-th column of B and the mean of y respectively. We solve this regression problem via the Cholesky decomposition of V, meaning we acquired

$$V = LL^T \tag{7}$$

, which allows us to solve the regression problem in $O(n^2)$ time by backward substitution since L is triangular.

While this has to be done once, update formula for the following regression problems are required to keep the complexity of following solves in $O(n^2)$ time, instead of $O(n^3)$, which the Cholesky decomposition would require. There exist algorithms to update L by a rank 1 update

$$L_{new}L_{new}^T = L_{old}L_{old}^T + huu^T \tag{8}$$

, where $u \in \mathbb{R}^{n\times 1}$ is the vector, whose outer product determines the update together with the multiplier $h \in \mathbb{R}$, in $O(n^2)$ time. However, the update formula for V are

$$V_{i,m+1} \mathrel{+}= \sum_{k=1}^{n} (B_{ik} - \overline{B_i}) B_{mk} * w1_k \tag{9}$$

$$V_{m+1,m+1} \mathrel{+}= \sum_{k=1}^{n} B_{mk}^2 * w2_k + o \tag{10}$$

, which means the last row *and* column of V have to be updated. The form of the update matrix is generally of rank 2. Furthermore, the question is generally how to decompose this matrix into the vector u required for the Cholesky update.