

# 1 Adaptive Metropolis algorithm

Suppose, that at time  $t - 1$  we have sampled the states  $X_0, X_1, \dots, X_{t-1}$ , where  $X_0$  is the initial state. Then a candidate point  $Y$  is sampled from the proposal distribution  $q_t(\cdot|X_0, \dots, X_{t-1})$ , which now may depend on the whole history  $(X_0, X_1, \dots, X_{t-1})$ . The candidate point  $Y$  is accepted with probability

$$\alpha(X_{t-1}, Y) = \min \left( 1, \frac{\pi(Y)}{\pi(X_{t-1})} \right), \quad (1)$$

in which case we set  $X_t = Y$ , and otherwise  $X_t = X_{t-1}$ . Observe that the chosen probability for the acceptance resembles the familiar acceptance probability of the Metropolis algorithm. However, here the choice for the acceptance probability is not based on symmetry (reversibility) conditions since these cannot be satisfied in our case – the corresponding stochastic chain is no longer Markovian. The exactness of the simulation is discussed separately.

The proposal distribution  $q_t(\cdot|X_0, \dots, X_{t-1})$  employed in the AM algorithm is a Gaussian distribution with mean at the current point  $X_{t-1}$  and covariance  $C_t = C_t(X_0, \dots, X_{t-1})$ .

The crucial thing regarding the adaptation is how the covariance of the proposal distribution depends on the history of the chain. In the AM algorithm this is solved setting  $C_t = s_d(X_0, \dots, X_{t-1}) + s_d \epsilon I_d$  after an initial period, where  $s_d$  is a parameter that depends only on dimension  $d$  and  $\epsilon > 0$  is a constant we may choose very small. Here  $I_d$  denotes the  $d$ -dimensional identity matrix. In order to start, we select an arbitrary, strictly positive definite, initial covariance  $C_0$ , according to our best prior knowledge (which may be quite poor). We select an index  $t_0 > 0$  for the length of an initial period and define

$$C_t = \begin{cases} C_0, & t \leq t_0, \\ s_d \mathbf{cov}(X_0, \dots, X_{t-1}) + s_d \epsilon I_d, & t > t_0. \end{cases} \quad (2)$$