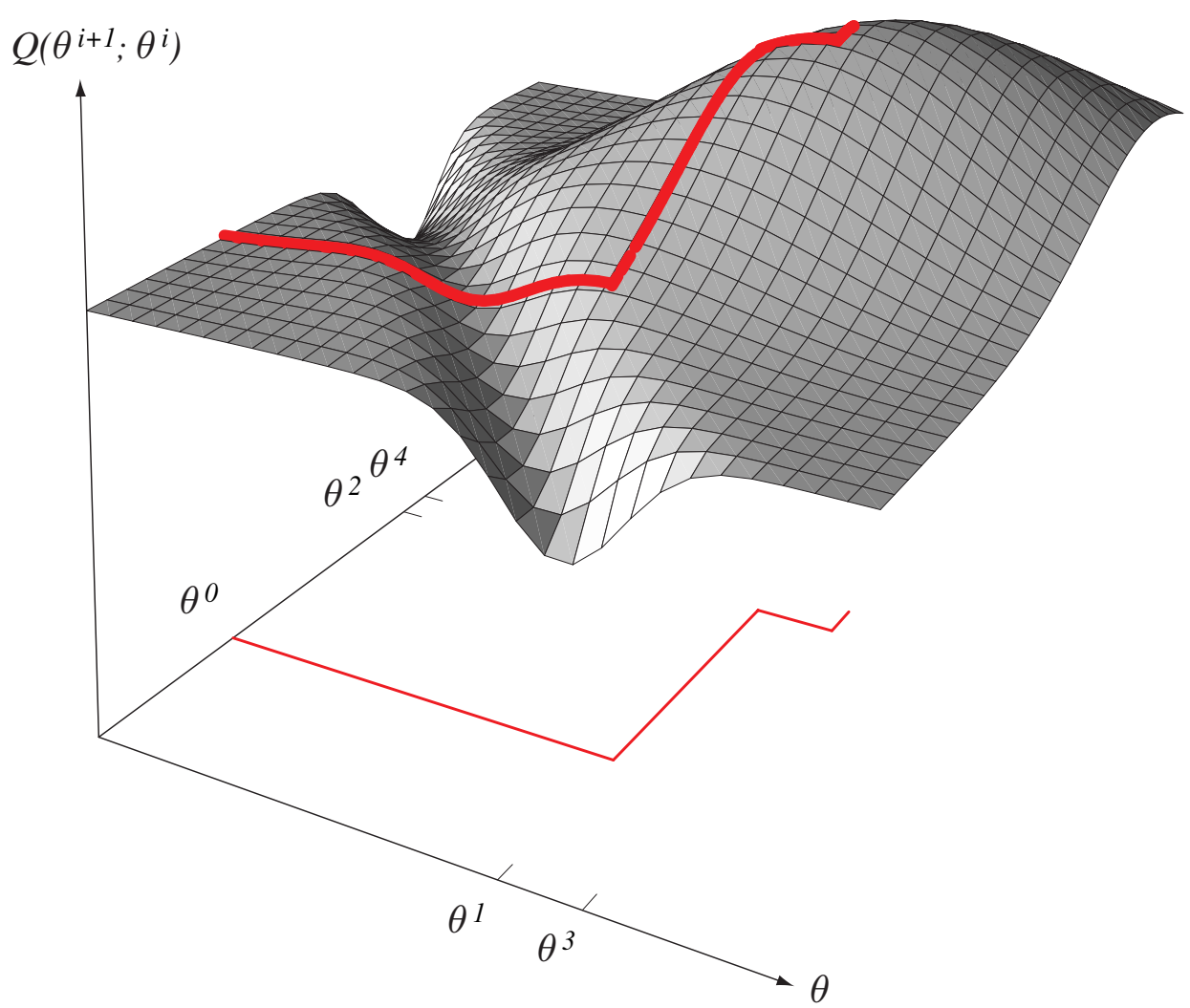

FIGURE 3.7. The search for the best model via the EM algorithm starts with some initial value of the model parameters, θ^0. Then, via the **M step**, the optimal θ^1 is found. Next, θ^1 is held constant and the value θ^2 is found that optimizes $Q(\cdot;\ \cdot)$. This process iterates until no value of θ can be found that will increase $Q(\cdot;\ \cdot)$. Note in particular that this is different from a gradient search. For example here θ^1 is the global optimum (given fixed θ^0), and would not necessarily have been found via gradient search. (In this illustration, $Q(\cdot;\ \cdot)$ is shown symmetric in its arguments; this need not be the case in general, however.) From: Richard O. Duda, Peter E. Hart, and David G. Stork, *Pattern Classification*. Copyright © 2001 by John Wiley & Sons, Inc.