9 Apr 1990 **Gradient Descent — Function Minimization**

Neil E Cotter

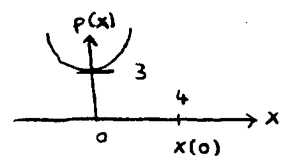
Gradient descent is a general purpose method for finding the minimum of a function. Unfortunately, we may only find a local minimum of the function instead of a global minimum.

ex: Find gradient descent algorithm that seeks out minimum of polynomial $p(x)$.

sol'n: $\Delta x = -\eta \dfrac{\partial p(x)}{\partial x}$ or $-\eta\, p'(x)$

ex: Given $p(x) = ax^2 + b$ $\quad p'(x) = 2ax$

$\Delta x = -\eta\, 2ax$ $\quad$ choose η small enough



Suppose $a = 2$ $\quad b = 3$.
Min is at $x = 0$.
Suppose $x(0) = 4$, $\eta = .1$

$x(1) = x(0) + \Delta x(0) = 4 + (-.1)\,2\cdot 2\cdot 4 = 4 - 1.6 = 2.4$

$x(2) = x(1) + \Delta x(1) = 2.4 + (-.1)\,2\cdot 2\cdot 2.4 = 2.4 - .96 = 1.$

$x(n+1) = x(n) + \Delta x(n) = x(n) - .4\,x(n) = .6\,x(n)$
$= x(0)\, .6^n$

We see that $x(n) \to 0$ as $n \to \infty$.

For this example we get convergence if $\eta\, 2a < 2$ or $\eta < \frac{1}{2}$.

We get divergence $(x(n) \to \infty)$ if $\eta > 1/2$.

Note: In neural network applications we minimize an error function. The message of this page is that we can minimize any differentiable function; it doesn't have to be an error function.