يكون حلًا عند $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ تمامًا عندما يكون للدالة g المعرّفة بالصيغة

$$g(x_1, x_2, \ldots, x_n) = \sum_{i=1}^{n} [f_i(x_1, x_2, \ldots, x_n)]^2$$

القيمة الصغرى صفرًا.

إن طريقة التناقص الأشد انحدارًا لإيجاد قيمة محلية صغرى لدالة ما من $\mathbb{R}^n$ إلى $\mathbb{R}$ يمكن وصفها حدسًا كما يأتي:

1. أوجد قيمة g عند نقطة تقريب ابتدائية $\mathbf{x}^{(0)} = \left(x_1^{(0)}, x_2^{(0)}, \ldots, x_n^{(0)}\right)^t$.
2. حدِّد اتجاهًا من $\mathbf{x}^{(0)}$ يسهم في تقليل قيمة g.
3. تحرك مسافة مناسبة في هذا الاتجاه. وعبِّر عن القيمة الجديدة بالرمز $\mathbf{x}^{(1)}$.
4. كرِّر العمليات من 1 إلى 3 باستخدام $\mathbf{x}^{(1)}$ بدلًا من $\mathbf{x}^{(0)}$.

إن جذر الكلمة اتجاه gradient هو الكلمة اللاتينية gradi التي تعني "تسير". بهذا المعنى يكون gradient لسطح ما هو معدل سرعة سيره إلى أعلى التلة.

قبل وصف كيفية اختيار الاتجاه الصحيح والمسافة المناسبة الواجب تحركها في هذا الاتجاه، نحتاج إلى مراجعة بعض نتائج التفاضل والتكامل.

تنص مبرهنة القيمة القصوى (The Extreme Value Theorem) على أن للدالة بمتغير واحد القابلة للاشتقاق قيمةً صغرى نسبية، عندما تكون المشتقة صفرًا فقط. لتعميم نتيجة الدوال هذه على متغيرات متعددة، نحتاج إلى تعريف الآتي

تعريف 9.10 للدالة $g: \mathbb{R}^n \to \mathbb{R}$، نعبِّر عن اتجاه g عند $\mathbf{x} = (x_1, x_2, \ldots, x_n)^t$ بالرمز $\nabla g(\mathbf{x})$ ونعرِّفه بالصيغة

$$\nabla g(\mathbf{x}) = \left(\frac{\partial g}{\partial x_1}(\mathbf{x}), \frac{\partial g}{\partial x_2}(\mathbf{x}), \ldots, \frac{\partial g}{\partial x_n}(\mathbf{x})\right)^t \qquad \blacksquare$$

إن الاتجاه gradient للدالة متعددة المتغيرات يشبه مشتقة الدالة وحيدة المتغير. بمعنى أن الدالة المتعددة المتغيرات القابلة للاشتقاق يمكن أن تملك قيمة صغرى نسبية عند x فقط إذا كان اتجاهها عند x هو المتجه الأصفر. يوجد للاتجاه خاصية أخرى مهمة مرتبطة بتصغير دوال متعددة المتغيرات. ليكن $\mathbf{v} = (v_1, v_2, \ldots, v_n)^t$ متجه الباب في $\mathbb{R}^n$

أي أن

$$\|\mathbf{v}\|_2^2 = \sum_{i=1}^{n} v_i^2 = 1$$

تعرف المشتقة المتجهة (directional derivative) للدالة g على x في اتجاه v بالصيغة

$$D_{\mathbf{v}} g(\mathbf{x}) = \lim_{h \to 0} \frac{1}{h}[g(\mathbf{x} + h\mathbf{v}) - g(\mathbf{x})] = \mathbf{v}^t \cdot \nabla g(\mathbf{x})$$

إن المشتقة المتجهة للدالة g على x في اتجاه v تقيس التغير في قيمة الدالة g. وبالنسبة إلى التغير في المتغير في اتجاه v هناك نتيجة في حساب التفاضل والتكامل للدوال متعددة المتغيرات تنص على أنه: عندما يكون g قابلًا للاشتقاق فإن الاتجاه الذي ينتج القيمة الكبرى للمشتقة المتجهة يحدث عندما يُختار v ليكون موازيًا لـ $\nabla g(\mathbf{x})$. على أن $\nabla g(\mathbf{x}) \neq \mathbf{0}$. ونتيجة لذلك يكون الاتجاه ذو النقص الأكبر في قيمة g عند x هو الاتجاه المعطى بـ $-\nabla g(\mathbf{x})$ (انظر شكل (3.10) لتوضيح حالة g دالة بمتغيرين.)

إن الغرض هو تقليل $g(\mathbf{x})$ إلى القيمة الصغرى وهي الصفر. ولذلك فإن الاختيار المناسب للمتجه $\mathbf{x}^{(1)}$ هو