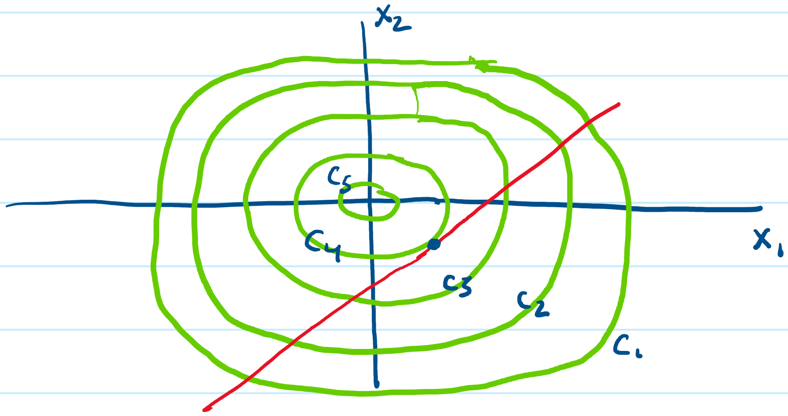
Lagrange Multipliers

Monday, August 15, 2022 2:52 PM

Problem: we wish to optimize $f(x)$ subject to the constraint that $g(x) = 0$.



— : contours of $f(\vec{x})$

— : $g(\vec{x}) = 0$

Note $\nabla f(\vec{x}_0)$ is the direction of steepest ascent at $\vec{x}_0$. We can prove this as follows.

$$\nabla f(\vec{x}) = \left[\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right]^T$$

The rate of change of f in the direction of $\vec{v}$ is

$$\left[v_1 \frac{\partial f}{\partial x_1} \quad v_2 \frac{\partial f}{\partial x_2} \quad \cdots \quad v_n \frac{\partial f}{\partial x_n} \right]^T = \nabla f \cdot \vec{v}.$$

$\nabla f \cdot \vec{v} = \|\nabla f\| \|\vec{v}\| \cos(\theta)$ where θ is the angle between ∇f and $\vec{v}$. $\underset{\theta}{\operatorname{argmax}} \cos\theta = 0$, So $\underset{\theta}{\operatorname{argmax}} \; \nabla f \cdot \vec{v} = 0.$

So the max rate of change of f is in the direction ∇f.

end note

We seek to minimize $f(\vec{x})$ among the points where $g(\vec{x}) = 0$. Note that when $(\nabla f)(\vec{x}_0) \parallel (\nabla g)(\vec{x}_0)$, moving along the constraint will not change $f(\vec{x})$, so we are at a local minimum or maximum. Lagrange multipliers allow us to find these points, which can be tested to determine the global maximum and minimum of f subject to constraint $g(\vec{x}) = 0$.

$(\nabla f)(\vec{x}_0) \parallel (\nabla g)(\vec{x}_0)$ means there is some λ such that $(\nabla f)(\vec{x}_0) = -\lambda (\nabla g)(\vec{x}_0)$.

Finding when

$\nabla f + \lambda \nabla g = 0$ for some λ will yield potential maxima and minima we desire.

The Lagrangian of a constrained optimization problem is defined as $\mathcal{L}(\vec{x}, \lambda) = f(\vec{x}) + \lambda g(\vec{x})$.

From above, we see we seek $\vec{x}^*$ and λ^* such that $\nabla \mathcal{L}(\vec{x}^*, \lambda^*) = 0$.

λ^* is found using the constraint(s).

Example: $f(\vec{x}) = \vec{x}^T \vec{x}$

$g(\vec{x}) = \vec{x}^T \vec{a} - b = 0$

$$\mathcal{L}(\vec{x}, \lambda) = \vec{x}^T \vec{x} + \lambda (\vec{x}^T \vec{a} - b)$$

$$\nabla \mathcal{L}(\vec{x}, \lambda) = 2\vec{x}^* + \lambda^* \vec{a} = 0$$

$$\vec{x}^* = -\frac{\lambda^* \vec{a}}{2}$$

$$\vec{x}^{*T} \vec{a} - b = 0$$

$$-\frac{\lambda^* \vec{a}^T \vec{a}}{2} - b = 0$$

$$\lambda^* = -\frac{2b}{\vec{a}^T \vec{a}}$$

$$\boxed{\vec{x}^* = -\frac{b}{\vec{a}^T \vec{a}} \vec{a} \qquad f(\vec{x}^*) = \frac{b^2}{\vec{a}^T \vec{a}}}$$