Contents

1 Optimisation

1.1 Introduction to optimisation

1.1.1 Introduction to unconstrained optimisation

1.1.1.1 Goals

We want to identify either the maximum or the minimum.

There exist local minima and global minima.

1.1.1.2 Optimising through limits

If we are looking to minimise a function, and the limits are ∞ or $-\infty$ then we can optimise by taking large or small values.

We can examine this for each variable.

This also applies for maximising a function.

1.1.1.3 Optimisation through stationary points

Stationary points of a function are points where marginal changes do not have an impact on the value of the function. As a result they are either local maxima or minima.

1.1.1.4 Optimisation through algorithms

If we cannot identify stationary points easily, we can instead use algorithms to identify optima.

1.1.1.5 Stationary points of strictly concave and convex functions

If a function is strictly concave it will only have one stationary point, a local, and global, maxima.

If a function is strictly convex it will only have one stationary point, a local, and global, minima.

1.1.2 Local optima

1.1.3 Optimising convex functions

1.1.4 Constrained optimisation

1.1.5 Analytic optimisation

1.1.5.1 Convex and concave functions

Convex functions only have one minimum, and concave functions have only one maximum.

If a function is not concave or convex, it may have multiple minima

If a function is convex, then there is only one critical point – the local minimum. We can identify this this by looking for critical points using first-order conditions.

Similarly, if a function is concave, then there is only one critical point – the local maximum.

We can identify whether a function is concave or convex by evaluating the Hessian matrix.

1.1.5.2 Evaluating multiple local optima

We can evaluate each of the local minima or maxima, and compare the sizes.

We can identify these by taking partial derivatives of the function in question and identifying where this function is equal to zero.

$u = f(x)$

$u_{x_i} = \frac{\delta f}{\delta x_i} = 0$

We can then solve this bundle of equations to find the stationary values of x.

After identifying the vector x for these points we can then determine whether or not the points are minima or maxima by examining the second derivative at these points. If it is positive it is a local minima, and therefore not an optimal point. Points beyond these will be higher, and may be higher than any local maxima.

1.2 Unconstrained optimisation of differentiable functions

1.2.1 Stationary points and first-order conditions

1.2.2 Stationary points and first-order conditions

1.2.3 Stationary points and first-order conditions

1.2.4 Hessian matrix

We can take a function and make a matrix of its second order partial derivatives. This is the Hessian matrix, and it describes the local curvature of the function.

If the function f has n parameters, the Hessian matrix is $n \times n$, and is defined as:

$H_{ij} = \frac{\delta^2 f}{\delta x_i \delta x_j}$

If the function is convex, then the Hessian matrix is positive semi-definite for all points, and vice versa.

If the function is concave, then the Hessian matrix is negative semi-definite for all points, and vice versa.

We can diagnose critical points by evaluating the Hessian matrix at those points.

If it is positive definite, it is a local minimum. If it is negative definite it is a local maximum. If there are both positive and negative eigenvalues it is a saddle point.

1.3 Linear optimisation

1.3.1 Single equality constraint

1.3.1.1 Constrained optimisation

Rather than maximise $f(x)$, we want to maximise $f(x)$ subject to $g(x) = 0$.

We write this, the Lagrangian, as:

$\mathcal{L}(x, \lambda) = f(x) - \sum_k^m \lambda_k [g_k(x) - c_k]$

We examine the stationary points for both vector x and λ. By including the latter we ensure that these points are consistent with the constraints.

1.3.1.2 Solving the Langrangian with one constraint

Our function is:

$\mathcal{L}(x, \lambda) = f(x) - \lambda[g(x) - c]$

The first-order conditions are:

$\mathcal{L}_\lambda = -[g(x) - c]$

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i}$

The solution is stationary so:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \lambda \frac{\delta g}{\delta x_i} = 0$

$\lambda \frac{\delta g}{\delta x_i} = \frac{\delta f}{\delta x_i}$

$\lambda = \frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}}$

Finally, we can use the following in practical applications.

$\frac{\frac{\delta f}{\delta x_i}}{\frac{\delta g}{\delta x_i}} = \frac{\frac{\delta f}{\delta x_j}}{\frac{\delta g}{\delta x_j}}$

1.3.2 Multiple equality constraints

1.3.2.1 Solving the Langrangian with many constraints

This time we have:

$\mathcal{L}_{x_i} = \frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = 0$

$\mathcal{L}_{x_j} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j} = 0$

$\frac{\delta f}{\delta x_i} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_i} = \frac{\delta f}{\delta x_j} - \sum_k^m \lambda_k \frac{\delta g_k}{\delta x_j}$

1.3.3 Inequality constraints

1.3.3.1 Lagrangians with inequality constraints

We can add constraints to an optimisation problem. These constraints can be equality constraints or inequality constraints. We can write constrained optimisation problem as:

Minimise $f(x)$ subject to

$g_i(x) \leq 0$ for $i = 1, \dots, m$

$h_i(x) = 0$ for $i = 1, \dots, p$

We write the Lagrangian as:

$\mathcal{L}(x, \lambda, \nu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \nu_i h_i(x)$

If we try and solve this like a standard Lagrangian, then all of the inequality constraints will instead by equality constraints.

1.3.3.2 Affinity of the Lagrangian

The Lagrangian function is affine with respect to λ and ν.

$\mathcal{L}(x, \lambda, \nu) = f(x) + \sum_{i=1}^m \lambda_i g_i(x) + \sum_{i=1}^p \nu_i h_i(x)$

$\mathcal{L}_{\lambda_i}(x, \lambda, \nu) = g_i(x)$

$\mathcal{L}_{\nu_i}(x, \lambda, \nu) = h_i(x)$

As the partial differential is constant, the partial differential is an affine function.

1.3.4 Primal and dual problems

1.3.4.1 The primal problem

We already have this.

1.3.4.2 The dual problem

We can define the Lagrangian dual function:

$g(\lambda, \nu) = \inf_{x \in X} \mathcal{L}(x, \lambda, \nu)$

That is, we have a function which chooses the returns the value of the optimised Lagrangian, given the values of λ and ν.

This is an unconstrained function.

We can prove this function is concave (how?).

The infimum of a set of concave (and therefore also affine) functions is concave.

The supremum of a set of convex (and therefore also affine) functions is convex.

Given a function with inputs x, what values of x maximise the function?

We explore constrained and unconstrained optimisation. The former is where restrictions are placed on vector x, such as a budget constraint in economics.

1.3.4.3 The dual problem is concave

1.3.4.4 The duality gap

We refer to the optimal solution for the primary problem as p^*, and the optimal solution for the dual problem as d^*.

The duality gap is $p^* - d^*$.

1.3.5 Complementary slackness for linear optimisation

1.3.6 Farkas' lemma

We have matrix A and vector b.

Either:

- $Ax = b$; $x \geq 0$
- $A^T y \geq 0$; $b^T y < 0$

1.4 Quadratic optimisation

1.4.1 The quadratic optimisation problem

1.5 Constrainted non-linear optimisation

1.5.1 Weak duality theorem

The duality gap ($p^* - d^*$ is non-negative.

1.5.2 Lagrange multipliers

1.5.3 The dual problem for non-linear optimisation

1.5.4 The weak duality theorem

1.6 Sort

1.7 Constrained convex optimisation

1.7.1 Slater's condition

1.7.1.1 Strong duality

Strong duality is where the duality gap is 0.

1.7.1.2 Slater's condition

Slater's condition says that strong duality holds if there is an input where the inequality constraints are satisified strictly.

That is they are $g(x) < 0$, not $g(x) \leq 0$

This means that the conditions are slack.

This only applies if the problem is convex. That is, if Slater's condition holds, and the problem is convex, then strong duality holds.

1.7.2 The strong duality theorem

1.7.3 Karush-Kuhn-Tucker conditions

If our problem is non-convex, or if Slater's condition does not hold, how else can be find a solution?

A solution, p^* can satisify KKT conditions.

1.7.4 Unconstrained envelope theorem

Consider a function which takes two parameters:

$f(x, \alpha$

We want to choose x to maximise f, given α.

$V(\alpha) = \sup_{x \in X} f(x, \alpha)$

There is a subset of X where $f(x, \alpha) = V(\alpha)$.

$X^*(\alpha) = \{x \in X | f(x, \alpha) = V(\alpha)\}$

This means that $V(\alpha) = f(x^*, \alpha)$ for $x^* \in X^*$.

Let's assume that there is only one x^*.

$V(\alpha) = f(x^*, \alpha)$

What happens to the value function as we relax α?

$V_{\alpha_i}(\alpha) = f_{\alpha_i}(x^*(\alpha), \alpha)$.

$V_{\alpha_i}(\alpha) = f_x \frac{\delta x^*}{\delta \alpha} + f_{\alpha_i}$.

We know that $f_x = 0$ from first order conditions. So:

$V_{\alpha_i}(\alpha) = f_{\alpha_i}$.

That is, at the optimum, as the constant is relaxed, we can treat the x^* as fixed, as the first-order movement is 0.