Lagrange Multipliers and the Rayleigh Quotient

Our goal is to find the extremum of the function $f(x) : R^n \to R$ subject to the constraints $h_i(x) = 0$ for $i = 1 \ldots m$ ($h_i : R^n \to R$).

Theorem (Lagrange multipliers): If x^* is an extremum of f subject to the constraints $h_i(x) = 0$ there exist scalars $\lambda_1, \ldots, \lambda_m$ such that

$$\nabla f(x^*) + \sum_{i=1}^{m} \lambda_i \nabla h_i(x^*) = 0 \tag{1}$$

where ∇f is the gradient of f. In other words, if x^* is an extremum subject to the constraints then

$$\frac{\partial f(x^*)}{\partial x_j} + \sum_{i=1}^{m} \lambda_i \frac{\partial h_i(x^*)}{x_j} = 0 \text{ for } j = 1 \ldots n \tag{2}$$

$\lambda_1, \ldots, \lambda_m$ are called Lagrange multipliers.

Example: The Rayleigh Quotient

$$\max_x \frac{x^T A x}{x^T x} \tag{3}$$

where A is symmetric.

Notice that $\frac{x'^T A x'}{x'^T x'} = \frac{x^T A x}{x^T x}$ for $x' = cx$ and $c \neq 0 \in R$, therefore we will solve for x with a unit norm $\|x\|_2^2 = 1$.

$$\begin{aligned} &\max x^T A x \\ &\text{s.t. } x^T x = 1 \end{aligned} \tag{4}$$

The Lagrangian is

$$L(x) = x^T A x + \lambda(x^T x - 1) \tag{5}$$

Taking the derivative with respect to x:

$$\frac{\partial L(x)}{\partial x} = x^T(A + A^T) + 2\lambda x^T \tag{6}$$
$$\frac{\partial L(x)}{\partial \lambda} = x^T x - 1 \text{ (the original constraint)} \tag{7}$$
$$\frac{\partial L(x)}{\partial x} = 0 \Rightarrow \tag{8}$$
$$x^T(A + A^T) = -2\lambda x^T \Rightarrow \tag{9}$$
$$(A + A^T)x = -2\lambda x \Rightarrow \text{(A is symmetric)} \tag{10}$$
$$Ax = \tilde{\lambda} x \text{ where } (\tilde{\lambda} = -2\lambda) \tag{11}$$

Hence the maximum and the minimum are obtained for x an eigenvector of A (the Lagrange multipliers provide a necessary condition. The extremum is indeed obtained because $x^T Ax$ is a continuous function and the unit sphere is a compact set). For x an eigenvector of A with unit norm, $x^T Ax = x^T \lambda x = \lambda x^T x = \lambda$. Therefore the maximum is obtained at the eigenvector corresponding to the largest eigenvalue of A.

The Generalized Rayleigh Quotient is:

$$\max_x \frac{x^T Ax}{x^T Bx} \tag{12}$$

For A, B symmetric and positive definite. Again, to choose a certain solution we will constrain x:

$$\begin{aligned} &\max\nolimits_x x^T x \\ &\text{s.t. } x^T Bx = 1 \end{aligned} \tag{13}$$

We will solve the Generalized Rayleigh Quotient by reduction to the Rayleigh Quotient.

Define $B = D^T D$, $C = D^{-T} AD^{-1}$ and $y = Dx$. Notice that $C \in PSDN$.

$$\frac{x^T Ax}{x^T Bx} = \frac{x^T D^T D^{-T} AD^{-1} Dx}{x^T D^T Dx} = \frac{y^T Cy}{y^T y} \tag{14}$$

This is the Rayleigh Quotient with the symmetric matrix C and the unit vector y ($y^T y = x^T D^T Dx = x^T BX = 1$). The solution is the first eigenvector of C. Notice that the first eigenvalue of C and $B^{-1}A$ is the same (substitute $y = Dx$ in $D^{-T} AD^{-1} y = \lambda y$), but their first eigenvector is different.