Implicit models

Using Lagrange multipliers to find *all* critical points

Math 561 spring 2023. (slides from K.Kubjas, with permission)

Lagrange multipliers

- Recall that the method of Lagrange multipliers is used to solve the following constrained optimization problem:

$$\max f(x)$$

$$\text{subject to } g_i(x) = 0 \text{ for } i = 1,\ldots,k$$

- The Lagrangian of this optimization problem is

$$L(x,\lambda) = f(x) - \sum_{i=1}^{k} \lambda_i g_i(x).$$

- Example: $L(x,\lambda) = l(p \mid u) - \lambda_1(p_{11} + p_{12} + p_{21} + p_{22} - 1) - \lambda_2(p_{11}p_{22} - p_{12}p_{21})$

Lagrange multipliers

The constrained critical points of f are among the unconstrained critical points of L. Hence one has to solve

$$g_1 = 0, \ldots, g_k = 0,$$

$$\frac{\partial f}{\partial x_1} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x_1} = 0, \ldots, \frac{\partial f}{\partial x_m} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial x_r} = 0$$

Lagrange multipliers

The gradient of the log-likelihood function is $\begin{pmatrix} \frac{u_1}{p_1} & \cdots & \frac{u_r}{p_r} \end{pmatrix}$. Hence:

$$g_1 = 0, \ldots, g_s = 0,$$

$$\frac{u_1}{p_1} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial p_1} = 0, \ldots, \frac{u_r}{p_r} - \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial p_r} = 0$$

Lagrange multipliers

- Clearing the denominators gives a system of polynomial equations:

$$g_1 = 0, \dots, g_s = 0,$$

$$u_1 - p_1 \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial p_1} = 0, \dots, u_r - p_r \sum_{i=1}^{k} \lambda_i \frac{\partial g_i}{\partial p_r} = 0$$

- When clearing the denominators, one might introduce new solutions where one of the p_i is zero (but this happens only if one of u_i is zero)

Lagrange multipliers

- In the statistical setting, one constraint is $p_1 + \ldots + p_r = 1$. Set $g_0 = p_1 + \ldots + p_r - 1$.
- Then $u_1 - p_1 \sum_{i=0}^{k} \lambda_i \frac{\partial g_i}{\partial p_1} = 0, \ldots, u_r - p_r \sum_{i=0}^{k} \lambda_i \frac{\partial g_i}{\partial p_r} = 0$ is equivalent to u being in the row span of the augmented Jacobian matrix

$$J' = \begin{pmatrix} p_1 & p_2 & \cdots & p_r \\ p_1 \frac{\partial g_1}{\partial p_1} & p_2 \frac{\partial g_1}{\partial p_2} & \cdots & p_r \frac{\partial g_1}{\partial p_r} \\ \vdots & \vdots & \ddots & \vdots \\ p_1 \frac{\partial g_k}{\partial p_1} & p_2 \frac{\partial g_k}{\partial p_2} & \cdots & p_r \frac{\partial g_k}{\partial p_r} \end{pmatrix}.$$

Lagrange multipliers

- Example:
  $L(x, \lambda) = l(p \mid u) - \lambda_1(p_{11} + p_{12} + p_{21} + p_{22} - 1) - \lambda_2(p_{11}p_{22} - p_{12}p_{21})$

- $p \in V(I)$ is a critical point of $l(p \mid u)$ if u is in the row span of the matrix
  $$\begin{pmatrix} p_{11} & p_{12} & p_{21} & p_{22} \\ p_{11}p_{22} & -p_{12}p_{21} & -p_{12}p_{21} & p_{11}p_{22} \end{pmatrix}$$

Lagrange multipliers

- Consider the ideal I_l generated by: $g_1, \ldots, g_s$,
$$u_1 - p_1 \sum_{i=0}^{k} \lambda_i \frac{\partial g_i}{\partial p_1}, \ldots, u_r - p_r \sum_{i=0}^{k} \lambda_i \frac{\partial g_i}{\partial p_r}.$$
- Whether the variety of the ideal is finite, can be checked with the command dim(I_l): dim=0 means that the system has finitely many solutions.
- If there are finitely many solutions, then the number of solutions can be computed with degree(I_l).
- The solutions can be found for example with the solve command in Mathematica.