

# ADMM

Tom Kealy

August 6, 2014

## 1 Constrained Convex Optimisation

This is a method of constrained optimisation, which adds a Lagrange multiplier like term to the an unconstrained objective (which coincides with the constrained optimisation problem). For example say you are trying to solve the following problem:

$$\min_x f(x) \tag{1.0.1}$$

$$\text{s.t } Ax = b \tag{1.0.2}$$

where  $f$  is a function in  $\mathbb{R}$ ,  $A \in \mathbb{R}^{n \times m}$  is a matrix, and  $x \in X$ ,  $X$  is some convex set.

This problem has Lagrangian:

$$L(x, y) = f(x) + y^T (Ax - b) \tag{1.0.3}$$

where  $y$  is the Lagrange multiplier.

This implies the dual problem:

$$g(y) = \inf_x L(x, y) = -f^*(-A^T y) - b^T y \tag{1.0.4}$$

where  $f^*$  is the convex conjugate (Legendre-Fenchel) transform of  $f$ . The problem is now to maximise  $g$ .

The solution is:

$$x^* = \arg \min_x L(x, y^*) \tag{1.0.5}$$

where  $y^*$  is the soln of the dual problem.

Together these suggest a iterative algorithm:

$$x^{k+1} := \arg \min_x L \tag{1.0.6}$$

$$y^{k+1} := y^k + \alpha^k (Ax^{k+1} - b) \tag{1.0.7}$$

This is a form of gradient ascent, with  $\alpha$  as the step size and  $k$  is the iteration counter. Assuming that  $g$  is differentiable, the algorithm is monotone (i.e.  $g(y^{k+1}) > g(y^k)$ ).

This algorithm is easily extended to the case where the objective function is linearly separable. For problems with  $g$  not differentiable, the augmented Lagrangian is introduced as:

$$L_\rho(x, y) = f(x) + y^T(Ax - b) + \frac{\rho}{2}\|Ax - b\|_2^2 \quad (1.0.8)$$

where  $\rho$  is a penalty parameter. The associated dual function is  $g_\rho(y) = \inf_x L_\rho$ . By adding the penalty term,  $g_\rho$  is no differentiable, and the  $\rho \rightarrow 0$  regime corresponds to the solution of the original problem.

The algorithm in this case can be found by first minimising over  $x$  and then evaluating the resulting equality constraint residual, i.e.:

$$x^{k+1} := \arg \min_x L_\rho \quad (1.0.9)$$

$$y^{k+1} := y^k + \rho(Ax^{k+1} - b) \quad (1.0.10)$$

This algorithm is known as the method of multipliers for solving the problem (2.0.16). Note that even if  $f$  is separable, the augmented Lagrangian may not be.

## 2 ADMM

The alternating direction method of multipliers (ADMM), is an algorithm intended to blend the decomposability of the gradient ascent algorithm 1.0.7, with the more general method of multipliers 1.0.10. The algorithm solves problems of the form

$$\min_x f(x) + g(z) \quad (2.0.11)$$

$$\text{s.t } Ax + Bz = c \quad (2.0.12)$$

where  $f$  and  $g$  are assumed to be convex function with range in  $\mathbb{R}$ ,  $A \in \mathbb{R}^{p \times n}$  and  $B \in \mathbb{R}^{p \times m}$  are matrices (not assumed to have full rank), and  $c \in \mathbb{R}^p$ .

For example the Basis Pursuit problem

$$\min_x \|x\|_1 \quad (2.0.13)$$

$$\text{s.t } Ax = b \quad (2.0.14)$$

can be written in the ADMM form:

$$\min_x f(x) + \|z\|_1 \quad (2.0.15)$$

$$\text{s.t } x - z = 0 \quad (2.0.16)$$

where  $f$  is the indicator function of  $\{x \in \mathbb{R}^n : Ax = b\}$ .

Similarly the LASSO:

$$\text{minimise } \frac{1}{2} \|Ax - b\|_2^2 + \lambda \|x\|_1 \quad (2.0.17)$$

$\lambda > 0$ , can be solved by the ADMM with  $f = \frac{1}{2} \|Ax - b\|_2^2$  and  $g = \lambda \|x\|_1$ .  
The solution of (2.0.12) is:

$$x* = \inf\{f(x) + g(z) : Ax + Bz = c\} \quad (2.0.18)$$

As previously we form the Lagrangian:

$$L_\rho(x, y) = f(x) + g(z) + y^T (Ax + Bz - c) + \frac{\rho}{2} \|Ax + Bz - c\|_2^2 \quad (2.0.19)$$

ADMM consists of the iterations:

$$x^{k+1} := \arg \min_x L_\rho(x, z^k, y^k) \quad (2.0.20)$$

$$z^{k+1} := \arg \min_z L_\rho(x^{k+1}, z, y^k) \quad (2.0.21)$$

$$y^{k+1} := y^k + \rho (Ax^{k+1} + Bz^{k+1} - c) \quad (2.0.22)$$