

# Maths Problem Set 4

Tim Munday  
OSM Boot Camp 2018

July 9, 2018

## Ex 6.6

For a point to be a critical point, it must have first derivative of zero in all arguments.

Differentiating with respect to  $x$ , and separately with respect to  $y$ , and setting these FOCs equal to zero, gives us four critical points.

For a (strict) local maximum, the Hessian evaluated at the point in question must be negative definite, for a (strict) local minimum, positive definite, and a saddle point if the Hessian is neither. We determine whether the Hessian is positive or negative definite by considering the eigenvalues. If both eigenvalues are positive then the matrix is positive definite, vice versa for negative definite. One positive and one negative eigenvalue indicates a saddle point.

$$\begin{aligned} A : x = \frac{-1}{3}, y = 0 & \quad | \quad \lambda_1 = \frac{1}{3}, \lambda_2 = -3 \Rightarrow \textit{Saddle Point} \\ B : x = \frac{-1}{9}, y = \frac{-1}{12} & \quad | \quad \lambda_1 = -0.309, \lambda_2 = -1.08 \Rightarrow \textit{Maximum} \\ C : x = 0, y = \frac{-1}{4} & \quad | \quad \lambda_1 = 1, \lambda_2 = -1 \Rightarrow \textit{Saddle Point} \\ D : x = 0, y = 0 & \quad | \quad \lambda_1 = -2, \lambda_2 = 0.5 \Rightarrow \textit{Saddle Point} \end{aligned}$$

## Ex. 6.11

The Newton method's first iteration gives:

$$x_1 = x_0 - \frac{f'(x_0)}{f''(x_0)}$$

Where  $f$  is a quadratic approximation (in this case the actual function).  
Evaluating this:

$$x_1 = \frac{-b}{2a}x_0$$

For  $x_1$  to be a minimiser (we know it is unique as the function itself is quadratic), we need the first derivative to be zero, and the second to be positive.

$$\begin{aligned} f'(x_1) &= 2ax_0 + b \\ &= -b + b \\ &= 0 \end{aligned}$$

$$f''(x_1) = 2a > 0$$

Therefore we have found the unique minimiser in one iteration. We could have expected this, because the Newton method approximates using a quadratic function, so the approx of a quadratic function is the function itself.

### Ex 7.1

$H - Conv(S)$  is convex if:

$$\lambda H_1 + (1 - \lambda)H_2 \in H$$

Expanding out, noting that  $H_i$  is a linear combination of the members of  $V$ , i.e. the  $x$ 's by definition. We have:

$$\begin{aligned} \lambda H_1 + (1 - \lambda)H_2 &= \lambda(\theta_1 x_1 + \theta_2 x_2 \dots + \theta_k x_k) + (1 - \lambda)(\rho_1 x_1 + \dots + \rho_k x_k) \\ &= (\lambda \theta_i + (1 - \lambda) \rho_i) x_i \in H \end{aligned}$$

Where the final line follows from the fact that the linear weighting of two series that sum to one will also sum to one, and as such we are essentially just applying a different weighting to the  $x$ 's, so they must be in  $H$  by definition of the convex hull.

### Ex 7.2

(i) A Hyperplane is a set of the form:

$$P = \{x \in V \mid \langle a, x \rangle = b\}$$

Taking two points in the hyper plane:  $X$  and  $Y$  and taking a linear combination of them, satisfies the first condition, since they are both in  $V$ , and the second is also satisfied:

$$\begin{aligned} \lambda \langle a, X \rangle + (1 - \lambda) \langle a, Y \rangle &= b \\ \lambda b + (1 - \lambda)b &= b \end{aligned}$$

(ii) We can follow the same procedure for a half-space:

$$H = \{x \in V | \langle a, x \rangle \leq b\}$$

Taking two points in the hyper plane:  $X$  and  $Y$  and taking a linear combination of them, satisfies the first condition, since they are both in  $V$ , and the second is also satisfied:

$$\begin{aligned} \lambda \langle a, X \rangle + (1 - \lambda) \langle a, Y \rangle &\leq b \\ \lambda p + (1 - \lambda)q &\leq b \\ \text{As } p &\leq b \\ \text{and } q &\leq b \end{aligned}$$

## 7.4

(i)

$$\begin{aligned} \|x - y\|^2 &= \|x - p - (y - p)\|^2 \\ &= \langle (x - p) - (y - p), (x - p) - (y - p) \rangle \\ &= \langle x - p, x - p \rangle + \langle p - y, p - y \rangle + 2\langle x - p, p - y \rangle \\ &= \|x - p\|^2 + \|p - y\|^2 + 2\langle x - p, p - y \rangle \end{aligned}$$

(ii)

From the definition of a projection we know that  $2\langle x - p, p - y \rangle \geq 0$ . We also know that the norm of order two is always positive so  $\|p - y\|^2 \geq 0$ . So, using the expression above we have:

$$\|x - y\|^2 = \|x - p\|^2 + K$$

Where  $K$  is strictly positive since  $y \neq p$ . Thus we have (square rooting both sides):

$$\|x - y\| > \|x - p\|$$

(iii)

Using the equation we proved in (i), we can write:

$$\begin{aligned} \|x - z\|^2 &= \|x - p\|^2 + \|p - z\|^2 + 2\langle x - p, p - z \rangle \\ &= \|x - p\|^2 + \|p - \lambda y + (1 - \lambda)p\|^2 + 2\langle x - p, p - (\lambda y + (1 - \lambda)p) \rangle \\ &= \|x - p\|^2 + 2\lambda \langle x - p, p - y \rangle + \lambda^2 \|y - p\|^2 \end{aligned}$$

(iv)

From (ii) we know that:

$$\|x - z\|^2 > \|x - p\|^2$$

Then we can rearrange (7.15) to get:

$$0 \leq 2\langle x - p, p - y \rangle + \lambda \|y - p\|^2$$

Since we know that this must hold even when  $\lambda$  is 0, we have the result:

$$\langle x - p, p - y \rangle \geq 0$$

### Ex 7.13

Let's assume  $f$  is bounded by  $M$ . That is  $f(x) \leq M \forall x$ . Take two points  $x_1$  and  $x_2$  in the domain of  $f$ . If  $f(x_1)$  and  $f(x_2)$  are not equal, then the line through them must cross  $y = M$  at some point. Since  $f$  is convex, we know that  $f(x)$  must lie on or above the line between  $f(x_1)$  and  $f(x_2)$  for all values of  $x, x_1, x_2$ . This is clearly not possible since the line will always intersect  $M$  unless the function is constant. Therefore the function is constant.

### Ex 7.20

If  $f$  is convex then we know that:

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \geq f(\lambda x_1 + (1 - \lambda)x_2)$$

If  $-f$  is convex then we know that:

$$\lambda f(x_1) + (1 - \lambda)f(x_2) \leq f(\lambda x_1 + (1 - \lambda)x_2)$$

Therefore:

$$\lambda f(x_1) + (1 - \lambda)f(x_2) = f(\lambda x_1 + (1 - \lambda)x_2)$$

So the function is affine.

### Ex. 7.21

Let  $x^* \in \mathbb{R}^n$  be a local minimizer of  $f$ . Then  $f(x^*) \leq f(x)$  for all  $x \in \mathcal{N}_r(x^*)$ , where  $\mathcal{N}_r(x^*)$  is an open ball around  $x^*$  of radius  $r > 0$ . Since  $\phi$  is monotonically increasing,  $\phi(f(x^*)) \leq \phi(f(x))$  for all  $x \in \mathcal{N}_r(x^*)$ . Thus,  $x^*$  is a local minimizer of  $\phi \circ f$ . Now let  $x^*$  be a local minimizer of  $\phi \circ f$ . Then  $\phi(f(x^*)) \leq \phi(f(x))$  for all  $x \in \mathcal{N}_r(x^*)$ , and since  $\phi$  is monotonically increasing, this implies that  $f(x^*) \leq f(x)$  for all  $x \in \mathcal{N}_r(x^*)$ . Thus,  $x^*$  is a local minimizer of  $f$ .