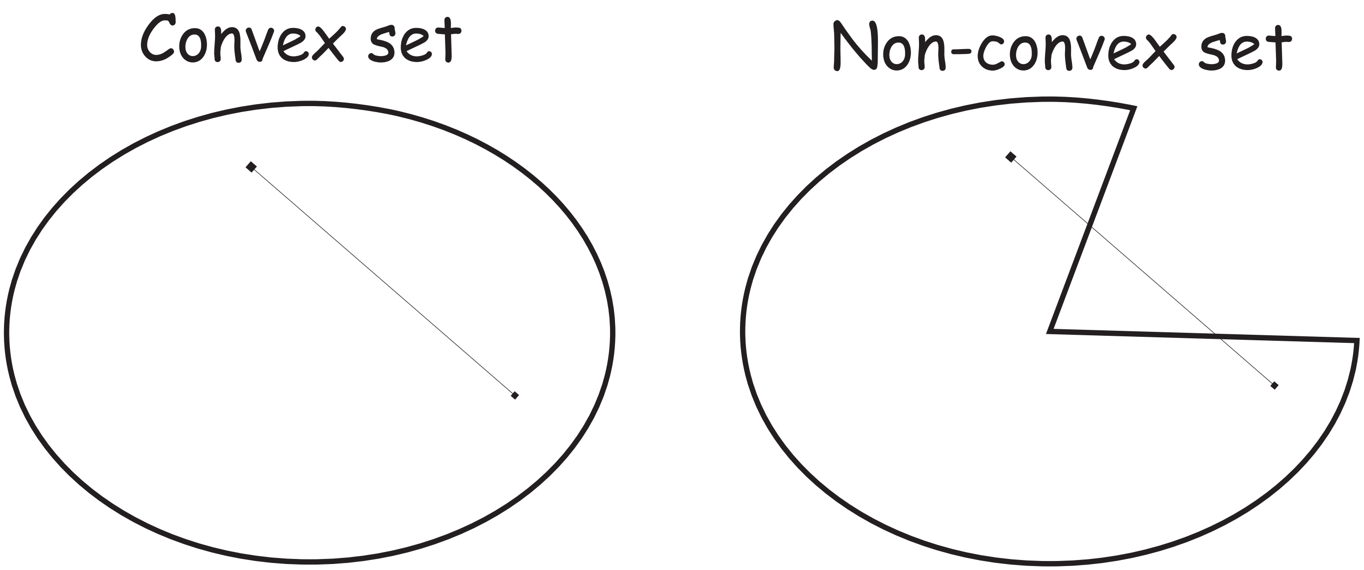
constrained convex optimization

virgil pavlu

convex set

a set X in a vector space is **convex** if for any $w_1, w_2 \in X$ and $\lambda \in [0, 1]$ we have $\lambda w_1 + (1 - \lambda) w_2 \in X$



convex function

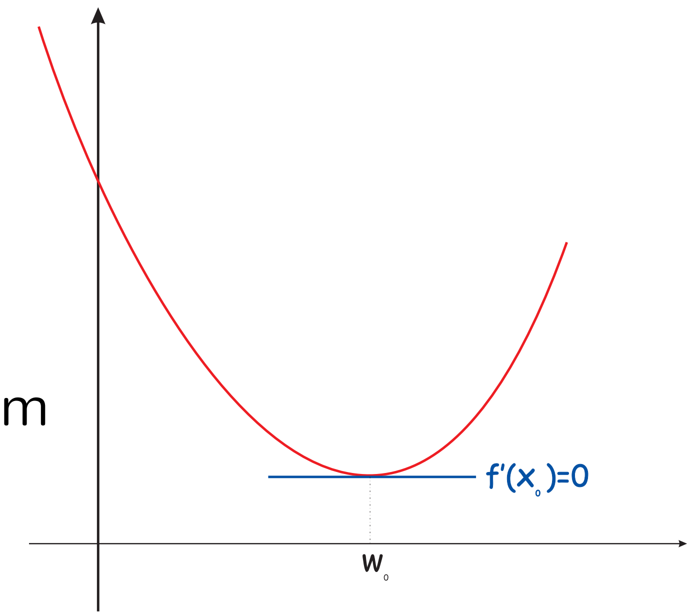
a function f is **convex(concave)** on $X \subseteq Dom(f)$ if for any $w_1, w_2 \in X$ and $\lambda \in [0, 1]$ we have
$f(\lambda w_1 + (1 - \lambda)w_2) \leq (\geq) \lambda f(w_1) + (1 - \lambda) f(w_2)$

if f is **strict convex** and **twice differentiable** on X then :

$\diamondsuit f\prime = \delta_w f(w)$ strict increasing

$\diamondsuit f\prime\prime \geq 0$

$\diamondsuit f\prime(x_0) = 0 \Leftrightarrow x_0$ is a global minimum



convex and differentiable

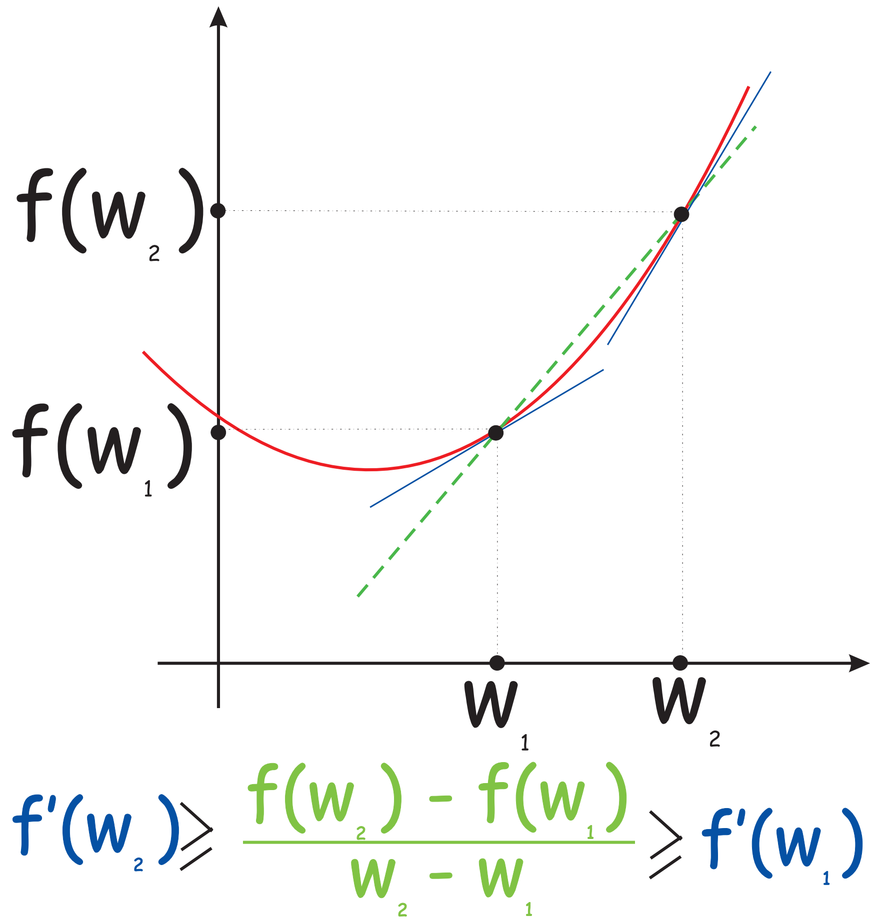
if f is convex and differentiable
then for any w_1,w_2 we have

1. $f'(w_2)(w_2 - w_1) \geq$
   $f(w_2) - f(w_1) \geq$
   $f'(w_1)(w_2 - w_1)$

2. $\exists w = \lambda w_1 + (1 - \lambda) w_2, \lambda \in [0, 1]$

such that

$$f(w_2) - f(w_1) = f'(w)(w_2 - w_1)$$



unconstrained optimization

one variable
interval cutting
Newton's Method

several variables
gradient descent
conjugate gradient descent

constrained optimization

given convex functions $f, g_1, g_2, ..., g_k, h_1, h_2, ..., h_m$ on convex set X, the problem

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i
$h_j(\mathbf{w}) = 0$,for all j

has as its solution a convex set. If f is strict convex the solution is unique (if exists)

we will assume all the good things one can imagine about functions $f, g_1, g_2, ..., g_k, h_1, h_2, ..., h_m$ like convexity, differentiability etc. That will still not be enough though....

equality constraints only

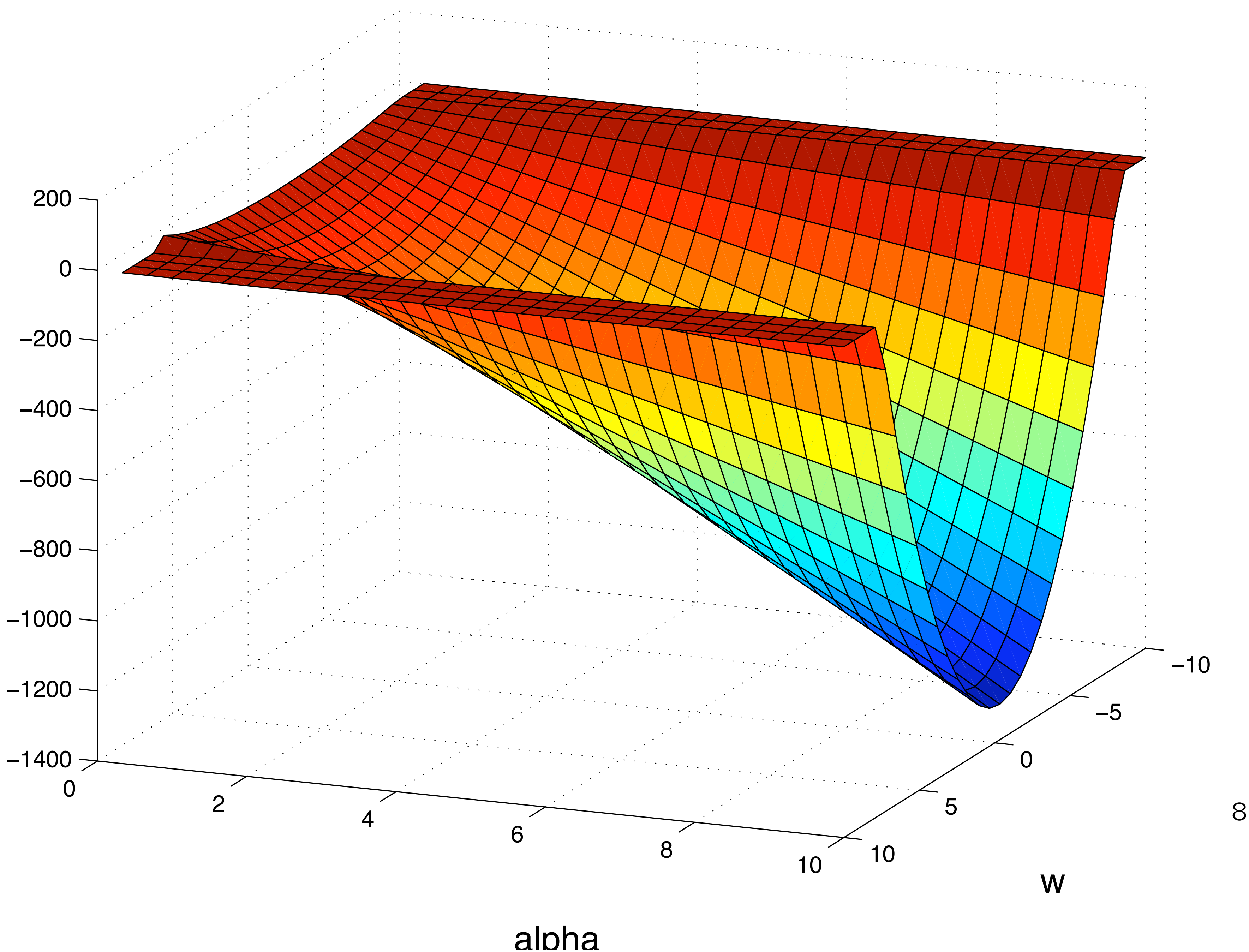
minimize $f(\mathbf{w})$
subject to $h_j(\mathbf{w}) = 0$,for all j

define the lagrangian function

$$L(\mathbf{w}, \beta) = f(\mathbf{w}) + \sum_j \beta_j h_j(\mathbf{w})$$

Lagrange theorem nec[essary] and suff[icient] conditions for a point $\tilde{\mathbf{w}}$ to be an optimum (ie a solution for the problem above) are the existence of $\tilde{\beta}$ such that

$$\delta_{\mathbf{w}} L(\tilde{\mathbf{w}}, \tilde{\beta}) = 0 \text{ ; } \delta_{\beta_j} L(\tilde{\mathbf{w}}, \tilde{\beta}) = 0$$


200
0
−200
−400
−600
−800
−1000
−1200
−1400
0
2
4
6
8
10
10
5
0
−5
−10
w
alpha
8

inequality constraints

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i
$h_j(\mathbf{w}) = 0$,for all j

we can rewrite every inequality constraint $h_j(\mathbf{x}) = 0$ as two inequalities $h_j(\mathbf{w}) \leq 0$ and $h_j(\mathbf{w}) \geq 0$. so the problem becomes

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i

Karush Kuhn Tucker theorem

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i
were g_i are **qualified constraints**

define the lagrangian function

$$L(\mathbf{w}, \alpha) = f(\mathbf{w}) + \sum_i \alpha_i g_i(\mathbf{w})$$

KKT theorem nec and suff conditions for a point $\tilde{\mathbf{w}}$ to be a solution for the optimization problem are the existence of $\tilde{\alpha}$ such that

$$\delta_{\mathbf{w}} L(\tilde{\mathbf{w}}, \tilde{\alpha}) = 0 \text{ ; } \widetilde{\alpha_i} g_i(\tilde{\mathbf{w}}) = 0$$

$$g_i(\tilde{\mathbf{w}}) \leq 0 \text{ ; } \widetilde{\alpha_i} \geq 0$$

KKT - sufficiency

Assume $(\tilde{\mathbf{w}}, \tilde{\alpha})$ satisfies KKT conditions

$\delta_{\mathbf{w}} L(\tilde{\mathbf{w}}, \tilde{\alpha}) = \delta_{\mathbf{w}} f(\tilde{\mathbf{w}}) + \sum_{i=1}^{k} \widetilde{\alpha_i} \delta_{\mathbf{w}} g_i(\tilde{\mathbf{w}}) = 0$
$\delta_{\alpha_i} L(\tilde{\mathbf{w}}, \tilde{\alpha}) = g_i(\tilde{\mathbf{w}}) \leq 0$
$\widetilde{\alpha_i} g_i(\tilde{\mathbf{w}}) = 0$; $\widetilde{\alpha_i} \geq 0$

Then
$f(\mathbf{w}) - f(\tilde{\mathbf{w}}) \geq (\delta_{\mathbf{w}} f(\tilde{\mathbf{w}}))^T (\mathbf{w} - \tilde{\mathbf{w}}) =$

$-\sum_{i=1}^{k} \widetilde{\alpha_i} (\delta_{\mathbf{w}} g_i(\tilde{\mathbf{w}}))^T (\mathbf{w} - \tilde{\mathbf{w}}) \geq -\sum_{i=1}^{k} \widetilde{\alpha_i} (g_i(\mathbf{w}) - g_i(\tilde{\mathbf{w}})) =$

$-\sum_{i=1}^{k} \widetilde{\alpha_i} g_i(\mathbf{w}) \geq 0$

so $\tilde{\mathbf{w}}$ is solution

saddle point

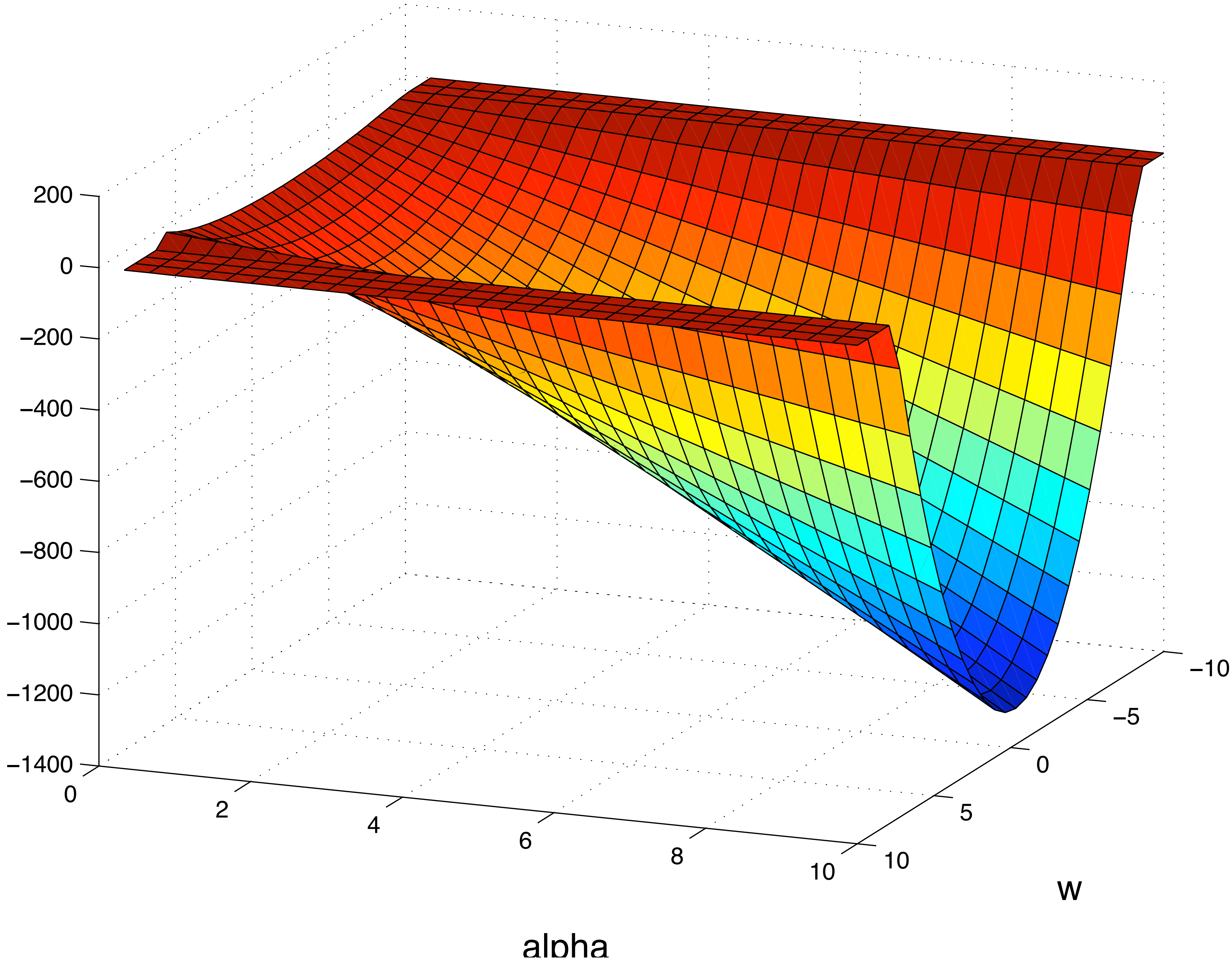
minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i

and the lagrangian function

$$L(\mathbf{w}, \alpha) = f(\mathbf{w}) + \sum_i \alpha_i g_i(\mathbf{w})$$

$(\tilde{\mathbf{w}}, \tilde{\alpha})$ with $\widetilde{\alpha_i} \geq 0$ is **saddle point** if $\forall (\mathbf{w}, \alpha), \alpha_i \geq 0$

$$L(\tilde{\mathbf{w}}, \alpha) \leq L(\tilde{\mathbf{w}}, \tilde{\alpha}) \leq L(\mathbf{w}, \tilde{\alpha})$$


200
0
−200
−400
−600
−800
−1000
−1200
−1400
0
2
4
6
8
10
alpha
−10
−5
0
5
10
w

saddle point - sufficiency

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i

lagrangian function $L(\mathbf{w}, \alpha) = f(\mathbf{w}) + \sum_i \alpha_i g_i(\mathbf{w})$
$(\tilde{\mathbf{w}}, \tilde{\alpha})$ is saddle point
$\forall(\mathbf{w}, \alpha), \alpha_i \geq 0 : L(\tilde{\mathbf{w}}, \alpha) \leq L(\tilde{\mathbf{w}}, \tilde{\alpha}) \leq L(\mathbf{w}, \tilde{\alpha})$

then

1. $\tilde{\mathbf{w}}$ is solution to optimization problem
2. $\widetilde{\alpha_i} g_i(\tilde{\mathbf{w}}) = 0$ for all i

saddle point - necessity

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$,for all i
were g_i are **qualified constraints**

lagrangian function $L(\mathbf{w}, \alpha) = f(\mathbf{w}) + \sum_i \alpha_i g_i(\mathbf{w})$
$\tilde{\mathbf{w}}$ is solution to optimization problem

then

$\exists \tilde{\alpha} \geq 0$ such that $(\tilde{\mathbf{w}}, \tilde{\alpha})$ is saddle point
$\forall (\mathbf{w}, \alpha), \alpha_i \geq 0 : \; L(\tilde{\mathbf{w}}, \alpha) \leq L(\tilde{\mathbf{w}}, \tilde{\alpha}) \leq L(\mathbf{w}, \tilde{\alpha})$

constraint qualifications

minimize $f(\mathbf{w})$
subject to $g_i(\mathbf{w}) \leq 0$, for all i

let Υ be the feasible region $\Upsilon = \{\mathbf{w} | g_i(\mathbf{w}) \leq 0 \ \forall i\}$

then the following additional conditions for functions g_i
are connected $(i) \Leftrightarrow (ii)$ and $(iii) \Rightarrow (i)$:

(i) there exists $w \in \Upsilon$ such that $g_i(\mathbf{w}) \leq 0 \ \forall i$
(ii) for all nonzero $\alpha \in [0, 1)^k \ \exists w \in \Upsilon$ such that $\alpha_i g_i(\mathbf{w}) = 0 \ \forall i$
(iii) the feasible region Υ contains at least two distinct elements, and $\exists w \in \Upsilon$ such that all g_i are are strictly convex at w w.r.t. Υ

KKT-gap

Assume $\tilde{\mathbf{w}}$ is the solution for optimization problem. Then for any $(\mathbf{w}, \alpha)$ with $\alpha \geq 0$ and satisfying

$$\delta_{\mathbf{w}} L(\mathbf{w}, \alpha) = 0 \text{ ; } \delta_{\alpha_i} L(\mathbf{w}, \alpha) \geq 0$$

we have

$$f(\mathbf{w}) \geq f(\tilde{\mathbf{w}}) \geq f(\mathbf{w}) + \sum_{i=1}^{k} \alpha_i g_i(\mathbf{w})$$

duality

$$f(\mathbf{w}) \geq f(\widetilde{\mathbf{w}}) \geq f(\mathbf{w}) + \sum_{i=1}^{k} \alpha_i g_i(\mathbf{w})$$

dual maximization problem :
maximize $L(\mathbf{w}, \alpha) = f(\mathbf{w}) + \sum_{i=1}^{k} \alpha_i g_i(\mathbf{w})$
subject to $\alpha \geq 0$ **;** $\delta_{\mathbf{w}} L(\mathbf{w}, \alpha) = 0$

OR
set $\theta(\alpha) = \inf_{\mathbf{w}} L(\mathbf{w}, \alpha)$
maximize $\theta(\alpha)$
subject to $\alpha \geq 0$

the primal and dual problems have the same objective value iff the gap can be vanished