

training functions
train gd

train gdm
adds .mc

train gda
adds .lr_inc
.lr_dec

train gdx
adds .mc to train gda

train rvp

train cgf
uses .beta

train cgp

train cgb

use a restart

and the search

srchcha = Charalambous search

train scg
uses .sigma

gd = gradient descent

m = momentum

a = adaptive (uses gd)

x = momentum

rvp = resilient backpropagation

cg = conjugate gradient
f = Fletcher-Reeves

$$x_{k+1} = x_k + \alpha_k p_k \quad ; \quad p_0 = -\nabla \text{Error} = -g_0$$

$$p_k = -g_k + \beta_k p_{k-1} \quad \& \quad \beta_k = \frac{\|g_k\|^2}{\|g_{k-1}\|^2}$$

$\|\cdot\| = \text{norm}$

$$\|g_k\|^2 = g_k^T g_k$$

p = Polak-Ribière

$$\beta_k = \frac{\Delta g_{k-1}^T g_k}{\|g_{k-1}\|^2}$$

b = Powell-Beale

$$\text{if } |g_{k-1}^T g_k| \geq 0.2 \|g_k\|^2$$

A = scaled

scales the weights by sigma on the 2nd derivative

EE434

02/28/07

uses λ

λ "regulates" the Hessian

train bfgs

bfgs = Broyden, Fletcher,
Goldfarb, Shanno

uses srbac = backtracking search

train oss

oss = one step secant

uses trinegf with srbac

train lm

lm = Levenberg-Marquardt