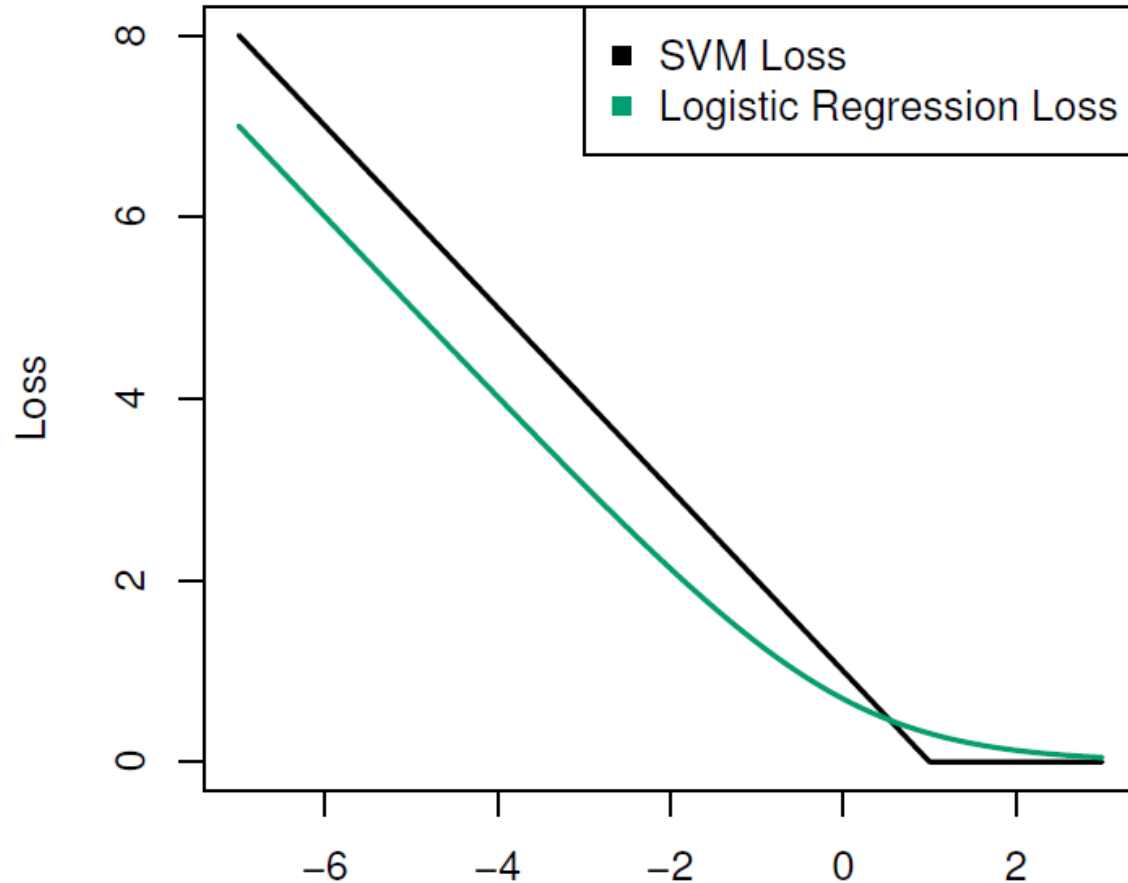


SVM naspram drugih modela

---

# Logistička regresija vs. SVM

- Veoma slična *loss* funkcija



# Logistička regresija vs. SVM

- Ukoliko je  $D$  veliko (u poređenju sa  $M$ ):
  - Npr. spam/ham klasifikacija gde imamo oko 10 000 obeležja (reči), a svega 1000 instanci (mailova)
  - Nemamo dovoljno podataka za fitovanje kompleksnih granica odluke
  - Koristiti logističku regresiju ili SVM bez kernela (*linear kernel*)
- Ukoliko je  $D$  malo, a  $N$  srednje veličine:
  - Npr.  $D \in [1, 1000]$ ,  $N \in [10, 10^4]$
  - Koristiti SVM sa RBF (Gausovim) kernelom
- Ukoliko je  $D$  malo, a  $N$  veliko:
  - Npr.  $D \in [1, 1000]$ ,  $N > 50\,000$
  - Koristiti logističku regresiju uz (manuelno) dodavanje obeležja ili SVM bez kernela (inače će biti veoma spor)
- Ako želite da estimirate verovatnoće, logistička regresija je bolji izbor

# Logistička regresija vs. SVM

- Logistička regresija i SVM bez kernela su obično veoma slični po pitanju performansi
  - Ali, u zavisnosti od detalja implementacije jedno može biti efikasnije od drugog
- Moć SVM-a dolazi do izražaja prilikom korišćenja kernela
- Ovo su samo generalne smernice za izbor klasifikacionog algoritma ali u praksi često nije sasvim jasno koji će raditi bolje
- Kao što smo već videli, algoritam jeste bitan, ali su često bitnije druge stvari – koliko podataka imamo, kakva su obeležja dostupna,...

# Drugi modeli i kerneli

- Može li se ova ideja sa kernelima primeniti i na druge klasifikacione algoritme (npr. logističku regresiju)?
- U principu, može, ali...
- Možemo konstruisati nova korisna obeležja
  - Ali dimenzionalnost problema postaje  $N$  (broj primera u trening skupu) – već za skup od 100 primera imamo 10 000 obeležja
  - Optimizacija algoritma veoma brzo postaje računarski neizvodljiva
- SVM ovo savladava pomoću kernel trika
  - Ne moramo eksplicitno graditi više-dimenzionu reprezentaciju i trenirati model u toj reprezentaciji
- Međutim, kernel trik (koji se izuzetno slaže sa SVM-om) je teško iskoristiti u kontekstu drugih optimizacionih funkcija

# Osobine SVM

- Fleksibilnost prilikom izbora funkcije sličnosti
- *Sparsness* rešenja kada se radi sa velikim skupom podataka
  - Granica odluke je određena samo podskupom podataka – vektorima potpore
- Velika dimenzionalnost nije velik problem
  - Kompleksnost ne zavisi od dimenzionalnosti problema
- Overfitting se kontroliše *soft margine* pristupom
- Lepa matematička svojstva
  - jednostavan konveksan optimizacioni problem za koji je garantovano da konvergira u globalni optimum

# SVM primene

- Kategorizacija teksta
- Prepoznavanje objekata na slikama
- Prepoznavanje rukom pisanih cifara
- Dijagnoza raka dojke
- Među algoritmima sa najboljim kvalitetom predviđanja u najrazličitijim domenima

# SVM danas

- Shujun et al. (2018): "Applications of Support Vector Machine (SVM) Learning in Cancer Genomics." <http://cgp.iijournals.org/content/15/1/41.full>
- *Zhang et al.* ICML 2017 Scaling Up Sparse Support Vector Machine by Simultaneous Feature and Sample Reduction  
<https://icml.cc/Conferences/2017/AcceptedPapers>
- Paiva et al. 2018. Supervised learning methods for pathological arterial pulse wave differentiation: A SVM and neural networks approach
- Kombinacije sa Deep learning pristupom
  - Tang (2013): „Deep learning using linear support vector machines“
  - Elleuchand Kherallah ( 2018): „An improved Arabic handwritten recognition system using deep support vector machines“
  - Zareapoor, et al. (2017): „Kernelized support vector machine with deep learning: an efficient approach for extreme multiclass dataset“