Stödvektormaskiner
Linjära hyperplan i Hilbertrum

Oscar Granlund

Kandidatavhandling i matematik
Fakulteten för naturvetenskaper och teknik
Åbo Akademi

9 november 2018

Stödvektormaskinen utvecklades under den senare halvan av 1900-talet i huvudsak av den ryske statistikern/datavetaren Vladimir Vapnik.

- Tog sin början år 1963 med en *linjär klassificerare* som endast gick att tillämpa på några problem.
- År 1992 presenterades en version som gick att tillämpa på alla problem.

Stödvektormaskinen går ut på att man skjuter in ett hyperplan mellan två klasser och använder hyperplanet för att klassificera nya observationer.

Parallellt med forskningen om stödvektormaskiner fann statistiker att en speciell typ av funktion, kärnor, kunde användas för att generalisera linjära algoritmer.

- Kärnorna föreslogs redan år 1964 för att generalisera en annan typ av linjär klassificerare.
- De användes även för att studera till exempel spline-modeller.
- År 1992 tillämpades kärnor på den ursprungliga stödvektormaskinmetoden.

Snart därefter (1995) tillämpades kärnor på den mera generaliserade algoritmen som presenterades 1992. Resultatet är en (tidsmässigt och resultatmässigt) effektiv *olinjär klassificerare* som ännu idag används.

Konvex optimering

Kvadratiska optimerignsproblem

De flesta av algoritmerna inom statistik och maskininlärning går att skriva om som konvexa optimeringsproblem. Ett optimeringsproblem är kvadratiskt och konvext om det går att skriva om på formen

$$\begin{aligned} \min \quad & \frac{1}{2}\mathbf{x}^\intercal \mathbf{P}\mathbf{x} + \mathbf{q}^\intercal \mathbf{x} + r \\ \text{så att} \quad & g_i(\mathbf{x}) \leq 0, \quad i = 1,\ \ldots,\ n \\ & h_i(\mathbf{x}) = 0, \quad i = 1,\ \ldots,\ m, \end{aligned}$$

där $\mathbf{P}$ är en positivt semidefinit matris, $g_i(\mathbf{x})$ är högst en kvadratisk funktion och alla krav g_i och h_i är satisfierbara samtidigt.

Ett konvext optimeringsproblem med olikhetskrav och likhetskrav kan lösas genom att man hittar alla extrempunkter (det borde bara finnas en). Detta kan göras med Lagrangemultiplikatorer. För ett kvadratiskt optimeringsproblem blir Lagrangefunktionen

$$L_P = f(\mathbf{x}) + \sum_{i=1}^{n} \lambda_i g_i(\mathbf{x}) + \sum_{i=1}^{m} v_i h_i(\mathbf{x})$$

där $f(\mathbf{x})$ är objektfunktionen och g_i, h_i är kraven.