■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks: ... For these additional metrics and the numerical classification example, we obtain the following values:

$$\text{acc} = (TP+TN)/(P+N) = (7+15)/(12+20) = \frac{22}{32} .$$

$$\text{sensitivity} = \text{sens} = TP/P = \frac{7}{12} .$$

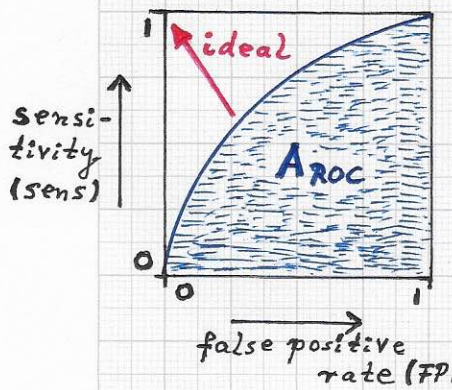
$$\text{specificity} = \text{spec} = TN/N = \frac{15}{20} .$$

$$\text{balanced accuracy} = \text{bacc} = \frac{1}{2}(\text{sens}+\text{spec}) = \frac{2}{3} .$$

(The numbers are those in the tables on the previous page.)

Metrics that are "complementary" to sensitivity and specificity are:

- FALSE NEGATIVE RATE = $FN/P = 1 - \text{sens}$.
- FALSE POSITIVE RATE = $FP/N = 1 - \text{spec}$.



Another metric considers the so-called receiver operating characteristic curve (ROC curve). This curve is often used to characterize the performance of a binary classification system. The curve is the graph of sensitivity vs. false positive rate, see figure on this page. Specifically, the performance metric is defined by the area A_{ROC}, the area under the curve, i.e.,

$$A_{ROC} = \int_{FPR=0}^{1} \text{sens}(FPR)\, d\,FPR .$$

Thus, the value $A_{ROC} = 1$ represents ideal performance.

...

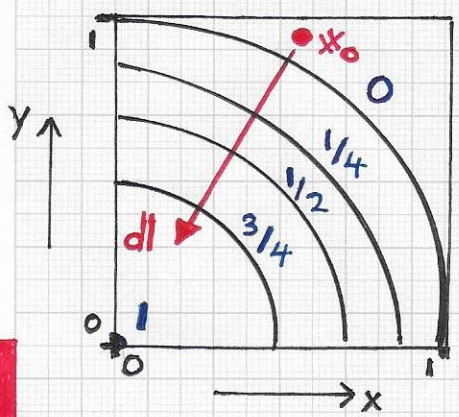
■ OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:...

In summary, MANY performance metrics exist or can be defined for a binary and multi-class classification system. A specific application and its overarching classification purpose and goal determine the metric to be used for system optimization. Once the metric is chosen, one must employ an optimization method – usually a combinatorial optimization method – to maximize the value of the metric.

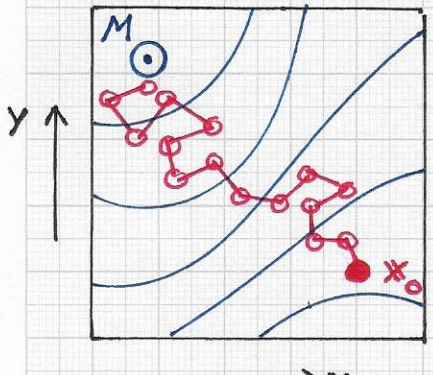
• SIMULATED ANNEALING – Combinatorial Optimization

Simulated annealing is a combinatorial optimization method one can use to optimize the values of the most relevant parameters of the classification system – subject to the specific performance metric chosen. First, we consider a simple analytical example. The figure (left) shows contours of the bivariate function $f(x,y) = 1 - x^2 - y^2$, where $x, y \in [0, 1]$. This function has its maximal value 1 at (0,0). Using a point x_0 as initial estimate, the steepest ascent method moves this point closer and closer to (0,0) by moving it in constant direction d. ...

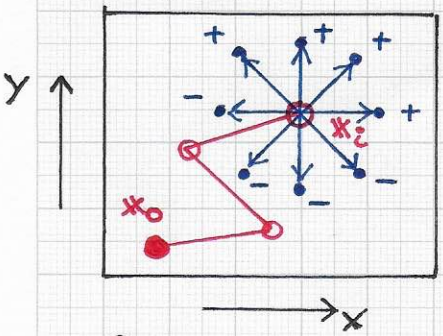


■ OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

• Laplacian eigenfunctions and neural networks:... How can one interpret the problem of determining this function's maximum as a discrete combinatorial (and stochastic) optimization problem?



Path from initial position x_0 towards maximum M.



Path can be continued in eight potential directions.

The two figures illustrate the case of finding the location M in the plane where a bivariate function is maximal. The top figure (left) indicates some of the function's contours (blue curves) and a polygon starting at position x_0 and ending at a position close to the maximum location M. Of special importance is the set of ordered vertices $\{v\}$ of this polygon (red). The vertex sequence does not seem to have an associated, corresponding sequence of function values such that these values only increase. This is indeed the case, and this BEHAVIOR CHARACTERIZES SIMULATED ANNEALING. The bottom figure is a sketch of a simulated annealing step. At vertex position x_i, one can pick from eight directions of possible progression. Here, four directions lead to points with a larger (smaller) function value '+' ('-'). ...

OBJECT AND MATERIAL EIGENFUNCTIONS - Cont'd.

- Laplacian eigenfunctions and neural networks...

In this example, simulated annealing always accepts a next vertex at position x_{i+1} when the function value increases, i.e., when $f(x_{i+1}) > f(x_i)$. SIMULATED ANNEALING ALSO ACCEPTS A NEXT VERTEX AT POSITION x_{i+1} EVEN WHEN THE FUNCTION VALUE DOES NOT INCREASE, I.E., WHEN $f(x_{i+1}) \leq f(x_i)$ — SUBJECT TO A PRE-DEFINED PROBABILITY FOR SUCH A POSITION TO BE ACCEPTED.

- **Note.** An intuitive explanation and rationale for this behavior of simulated annealing is the following scenario: "When climbing a hill with hillsides filled with large rocks, one does not want to end up on top of one of the rocks; the goal is to reach the peak of the hill. Thus, one must occasionally move downwards and go around a rock before moving upwards again." In other words, the purpose of using simulated annealing is the desire to maximize a function globally, in the presence of many local maxima. Generally, simulated annealing is viewed as a minimization method. Its name reflects the conceptual relationship to energy minimization in thermodynamics.

...

OBJECT AND MATERIAL EIGENFUNCTIONS – Cont'd.

• Laplacian eigenfunctions and neural networks:... When cooling down, many materials attempt to minimize their energy, often reaching a globally minimal energy state ("annealing" of metals). The underlying thermodynamical process causes a material to continually change – generally to bring down – its temperature / energy level. This process allows a material to also increase its temperature / energy with a certain probability; this probability goes down as the minimization process progresses. The high-level summary of the core part of a simulated annealing algorithm is provided here (left). In the context of globally maximizing a multivariate function $f(x_1, \ldots, x_n)$, one starts with an initial tuple / configuration $(x_1, x_2, \ldots, x_{n-1}, x_n)$; one subsequently changes this configuration – in a discrete, combinatorial setting – following the rules of simulated annealing. FOR OUR APPLICATION, MATERIAL CLASSIFICATION, WE MUST MAXIMIZE THE CHOSEN PERFORMANCE FUNCTION BY OPTIMIZING ITS MOST IMPORTANT PARAMETER VALUES.

- An initial configuration is given.
- As long as certain termination conditions are not satisfied:
  - i) Consider a configuration change.
  - ii) Would this change reduce energy?
    - a) Yes ⇒ Execute the change.
    - b) No ⇒ Execute the change with a certain probability.

Core of simulated annealing.