Measures of Fitness

- Normalized distance

- Baseline Score

- Correlation.

Matching mnt20 to mnt 10

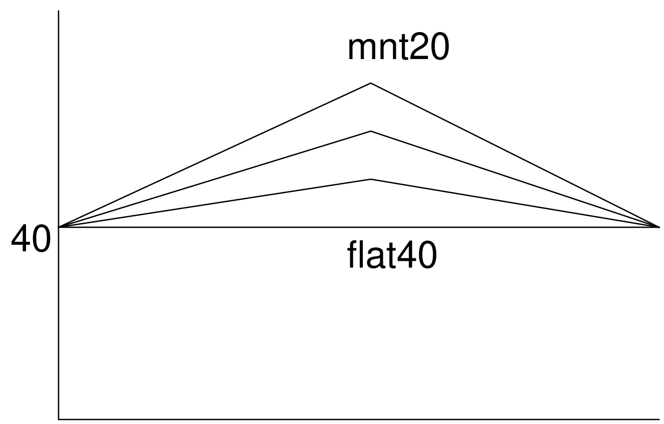
	0	1	2	3	4	5	6
6	90	50	70	110	130	90	70
5	90	30	50	90	90	70	70
4	80	20	40	60	70	60	80
3	60	20	20	50	60	70	100
2	30	10	30	70	90	90	110
1	10	10	40	90	120	130	140
0	0	20	60	120	160	180	180

DP Formulation:

Cumulative Distance

$$D(i,j) = d_{ij} + \min[D(i-1, j), D(i-1, j-1), D(i, j-1)]$$

Sample Problem:



DP Formulation

Stage Variable: Monotonicity (Time is used as the stage variable)

State Variable: Points on the grid

Decision Variable: Restrictions on transitions (critical for restricting search)

Possible restrictions:

- Monotonicity: $x_i >= x_{i-1}$, $y_i >= y_{i-1}$
- Continuity: $x_i - x_{i-1} <= 1$, $y_i - y_{i-1} <= 1$
- Warping Windows: $|x_k - y_k| <= w$
- Slope: Restrict slope
- Boundary Conditions: $x_1 = y_1 = 1$; $x_p = n$, $y_p = m$

DP Formulation.

First thing to formalize is the notion of distance:

dij = |si - tj|

dij = (di - tj)^2

$DTW(S,T) = min_W [Sum_k (d_{wk})]$

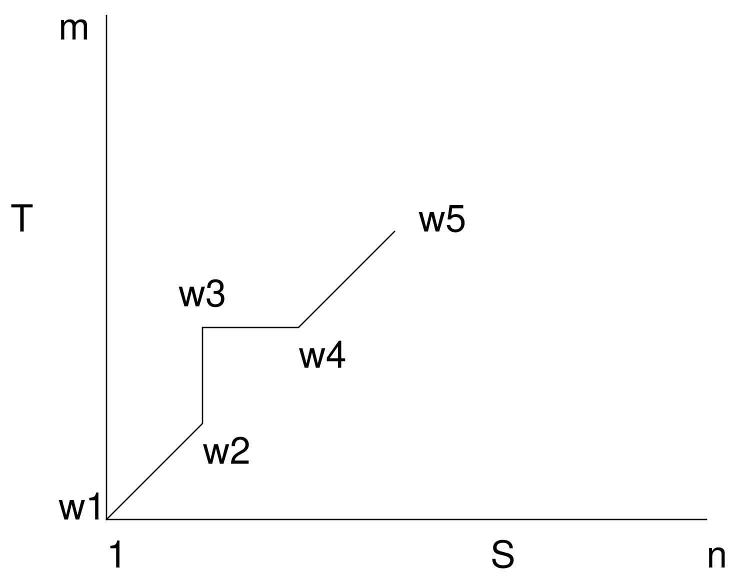
Warping:

Warping path maps / aligns elements of S to T such that a certain distance metric is minimized.

W = w1 w2 w3 ... wp

W corresponds to a sequence of grid points on a 2-D plane

Each wi corresponds to a tuple (i,j) indicating that si is aligned with tj. No warping => i = j



Dynamic Time Warping

Based on Dynamic Programming

DP attempts to find alignments that minimize a certain distance metric.

Time series is stretched/compressed for matching.

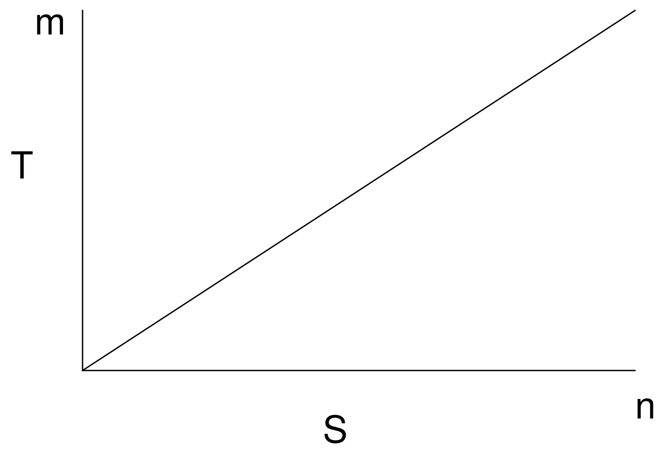
S = s1 s2 s3 ... sn
T = t1 t2 t3 ... tn

(i,j) represents the alignment of si and tj.

Dynamic Time Warping

Example: Speech Signals

In general, many applications with predetermined templates have this.



Matching string S with T.

Time series analysis: Looking for Known Templates

Consider the Gene Mapping Problem.

You have a string defined on the alphabet ATGC and you are matching proteins in this.

Complications:

Heuristics for Context

Gaps.

Other Applications:

Stock market data

Population demographics

Speech processing

(notion of fuzzy matching)