Modelling of Complex Real-World Systems
Part D. Applications in Computer Science

D1. Binary Classification

Ton Coolen
Department of Biophysics, Radboud University
module NWI-NM127, May 2021

Linear separability of data

Motivation: dimension mismatch and overfitting

two clinical outcomes (A,B),
4 patients, 60 gene expression levels ...

A : (100101001010010101010010001010111001001001001001001000011111)
A : (010001000010101001010101010010101000111100101001001010101000)
B : (001010001110101101100100100111001110010100101010101000101010)
B : (101011001010110010100100111100100101100111010111010001010010)

prognostic signature!

A : (100101001010010101010010001010111001001001001001001000011111)
A : (010001000010101001010101010010101000111100101001001010101000)
B : (001010001110101101100100100111001110010100101010101000101010)
B : (101011001010110010100100111100100101100111010111010001010010)

shuffle outcome labels ...

A : (100101001010010101010010001010111001001001001001001000011111)
B : (010001000010101001010101010010101000111100101001001010101000)
A : (001010001110101101100100100111001110010100101010101000101010)
B : (101011001010110010100100111100100101100111010111010001010010)

overfitting, no reproducibility ...

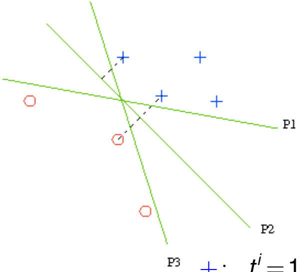
suppose we have data D on N patients,
pairs of measurement vectors + clinical outcome labels

$$D = \{(\mathbf{x}^1, t^1), \ldots, (\mathbf{x}^N, t^N)\}, \qquad \mathbf{x}^i \in \{-1,1\}^p, \quad t^i \in \{-1,1\}, \qquad p, N \gg 1$$

e.g. $\mathbf{x}^i$= gene expressions of patient i (each component on/off)
t^i= treatment response of i (yes/no)

- assumed model:

$$t(\mathbf{x}) = \begin{cases} 1 & \text{if } \sum_{\mu=1}^p \theta_\mu x_\mu > 0 \\ -1 & \text{if } \sum_{\mu=1}^p \theta_\mu x_\mu < 0 \end{cases}$$
$$= sgn\Big[\sum_{\mu=1}^p \theta_\mu x_\mu\Big]$$



$+$: $t^i = 1$
o: $t^i = -1$

- classification task:
find parameters $(\theta_1 \ldots, \theta_p)$ such that *for all* $i = 1 \ldots N$: $t^i = sgn\Big[\sum_{\mu=1}^p \theta_\mu x_\mu^i\Big]$

- data D explained perfectly by model with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_p)$ if

$$for\ all\ i = 1 \ldots N: \qquad t^i = sgn(\boldsymbol{\theta} \cdot \mathbf{x}^i), \quad i.e.\ \ t^i(\boldsymbol{\theta} \cdot \mathbf{x}^i) > 0$$

$$separating\ plane\ in\ input\ space: \qquad \boldsymbol{\theta} \cdot \mathbf{x} = 0$$

$$distance\ \Delta_i\ between\ \mathbf{x}^i\ and\ separating\ plane: \qquad d_i = t^i(\boldsymbol{\theta} \cdot \mathbf{x}^i)/|\boldsymbol{\theta}|$$

$$|\boldsymbol{\theta}|\ irrelevant,\ so\ choose\ |\boldsymbol{\theta}|^2 = p$$

- defn: *version space*

all $\boldsymbol{\theta}$ that solve above eqns with distances κ or larger

volume of version space:

$$V(\kappa) = \int \mathrm{d}\boldsymbol{\theta}\ \delta(\boldsymbol{\theta}^2 - p) \prod_{i=1}^{N} \theta\Big[\frac{t^i(\boldsymbol{\theta} \cdot \mathbf{x}^i)}{\sqrt{p}} - \kappa\Big]$$

(i) if $V(\kappa)$ large: many choices of $\boldsymbol{\theta}$ possible, easy problem
(ii) if $V(\kappa)$ small: few choices of $\boldsymbol{\theta}$ possible, hard problem
(iii) if $V(\kappa) = 0$: data not linearly separable

- high dimensional data: p large, $\alpha = N/p$
  $\alpha > 1$: fewer parameters than constraints
  $\alpha < 1$: more parameters than constraints

  $V(\kappa)$ scales exponentially with p,
  so define

$$F = \frac{1}{p}\log V(\kappa) \;=\; \frac{1}{p}\log\int \mathrm{d}\boldsymbol{\theta}\;\delta(\boldsymbol{\theta}^2 - p)\prod_{i=1}^{N}\theta\Big[\frac{t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)}{\sqrt{p}} - \kappa\Big]$$

  since $V(\kappa) = \mathrm{e}^{pF}$:

  $F = -\infty$: $V(\kappa) = 0$, no solutions $\boldsymbol{\theta}$ exist,
  data D not linearly separable

  $F =$ finite: $V(\kappa) > 0$, solutions $\boldsymbol{\theta}$ exist,
  data D linearly separable

Overfitting

parameters $\boldsymbol{\theta}$ define a 'rule' for computing answers t^i to questions $\mathbf{x}^i$, what if we choose *random data* D?

$$D=\{(\mathbf{x}^1,t^1),\ldots,(\mathbf{x}^N,t^N)\},\qquad \mathbf{x}^i\in\{-1,1\}^p,\ \ t^i\in\{-1,1\},\ \ \textit{fully random}$$

typical classification performance:

$$\overline{F} = \overline{\frac{1}{p}\log\int \mathrm{d}\boldsymbol{\theta}\ \delta(p-\boldsymbol{\theta}^2)\prod_{i=1}^{N}\theta\Big[\frac{t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)}{\sqrt{p}}-\kappa\Big]}$$

$$= \overline{\frac{1}{p}\log\int\frac{\mathrm{d}z}{2\pi}\mathrm{e}^{\mathrm{i}zp}\int\mathrm{d}\boldsymbol{\theta}\ \mathrm{e}^{-\mathrm{i}z\boldsymbol{\theta}^2}\prod_{i=1}^{N}\theta\Big[\frac{t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)}{\sqrt{p}}-\kappa\Big]}$$

introduce δ-functions,

$$1=\int\mathrm{d}y_i\ \delta\Big[y_i-\frac{t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)}{\sqrt{p}}\Big]=\int\frac{\mathrm{d}y_i\mathrm{d}\hat{y}_i}{2\pi}\ \mathrm{e}^{\mathrm{i}\hat{y}_iy_i-\mathrm{i}\hat{y}_it^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)/\sqrt{p}}$$

$$\overline{F}=\overline{\frac{1}{p}\log\int\frac{\mathrm{d}z\mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\mathrm{d}\boldsymbol{\theta}}{(2\pi)^{N+1}}\mathrm{e}^{\mathrm{i}zp+\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}-\mathrm{i}z\boldsymbol{\theta}^2}\Big(\prod_{i=1}^{N}\theta(y_i-\kappa)\mathrm{e}^{-\mathrm{i}\hat{y}_it^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)/\sqrt{p}}\Big)}$$

Gardner theory

large p, large N,
$N = \alpha p$:

$$\overline{F} = \lim_{p\to\infty} \frac{1}{p} \overline{\log \int \frac{\mathrm{d}z\mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\mathrm{d}\boldsymbol{\theta}}{(2\pi)^{N+1}} \mathrm{e}^{\mathrm{i}zp+\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}-\mathrm{i}z\boldsymbol{\theta}^2} \left(\prod_{i=1}^{N} \theta(y_i-\kappa)\mathrm{e}^{-\mathrm{i}\hat{y}_i t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)/\sqrt{p}} \right)}$$

- replica identity

  $\overline{\log Z} = \lim_{n\to 0} n^{-1} \log \overline{Z^n}$

$$\overline{F} = \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \overline{\left[\int \frac{\mathrm{d}z\mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\mathrm{d}\boldsymbol{\theta}}{(2\pi)^{N+1}} \mathrm{e}^{\mathrm{i}zp+\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}-\mathrm{i}z\boldsymbol{\theta}^2} \left(\prod_{i=1}^{N} \theta(y_i-\kappa)\mathrm{e}^{-\mathrm{i}\hat{y}_i t^i(\boldsymbol{\theta}\cdot\mathbf{x}^i)/\sqrt{p}} \right) \right]^n}$$

$$= \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \int \prod_{\alpha=1}^{n} \left[\frac{\mathrm{d}z^\alpha\mathrm{d}\mathbf{y}^\alpha\mathrm{d}\hat{\mathbf{y}}^\alpha\mathrm{d}\boldsymbol{\theta}^\alpha}{(2\pi)^{N+1}} \mathrm{e}^{\mathrm{i}pz^\alpha+\mathrm{i}\hat{\mathbf{y}}^\alpha\cdot\mathbf{y}^\alpha-\mathrm{i}z^\alpha(\boldsymbol{\theta}^\alpha)^2} \prod_{i=1}^{N} \theta[y_i^\alpha-\kappa] \right]$$
$$\times \overline{\mathrm{e}^{-\mathrm{i}\sum_{i=1}^{N}\sum_{\alpha=1}^{n} \hat{y}_i^\alpha t^i(\boldsymbol{\theta}^\alpha\cdot\mathbf{x}^i)/\sqrt{p}}}$$

- average over data D

$$
\begin{aligned}
\Xi &= \overline{\mathrm{e}^{-\mathrm{i}\sum_{i=1}^{N}\sum_{\alpha=1}^{n}\hat{y}_i^\alpha t^i(\boldsymbol{\theta}^\alpha\cdot\mathbf{x}^i)/\sqrt{p}}} = \overline{\mathrm{e}^{-\mathrm{i}\sum_{\mu=1}^{p}\sum_{i=1}^{N} t^i x_\mu^i \sum_{\alpha=1}^{n}\hat{y}_i^\alpha\theta_\mu^\alpha/\sqrt{p}}} \\
&= \prod_{\mu=1}^{p}\prod_{i=1}^{N}\overline{\mathrm{e}^{-\mathrm{i}x_\mu^i\sum_{\alpha=1}^{n}\hat{y}_i^\alpha\theta_\mu^\alpha/\sqrt{p}}} = \prod_{\mu=1}^{p}\prod_{i=1}^{N}\cos\left[\frac{1}{\sqrt{p}}\sum_{\alpha=1}^{n}\hat{y}_i^\alpha\theta_\mu^\alpha\right] \\
&= \prod_{\mu=1}^{p}\prod_{i=1}^{N}\left\{1-\frac{1}{2p}\Big(\sum_{\alpha=1}^{n}\hat{y}_i^\alpha\theta_\mu^\alpha\Big)^2+\mathcal{O}(\frac{1}{p^2})\right\} \\
&= \mathrm{e}^{-\frac{1}{2p}\sum_{\mu=1}^{p}\sum_{i=1}^{N}\sum_{\alpha,\beta=1}^{n}\hat{y}_i^\alpha\hat{y}_i^\beta\theta_\mu^\alpha\theta_\mu^\beta+\mathcal{O}(p^0)}
\end{aligned}
$$

$$
\begin{aligned}
\overline{F} &= \lim_{p\to\infty}\lim_{n\to 0}\frac{1}{pn}\log\int\prod_{\alpha=1}^{n}\left[\frac{\mathrm{d}z^\alpha\mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}^\alpha\mathrm{d}\boldsymbol{\theta}^\alpha}{(2\pi)^{N+1}}\mathrm{e}^{\mathrm{i}pz^\alpha+\mathrm{i}\hat{\mathbf{y}}^\alpha\cdot\mathbf{y}^\alpha-\mathrm{i}z(\boldsymbol{\theta}^\alpha)^2}\prod_{i=1}^{N}\theta(y_i^\alpha-\kappa)\right] \\
&\qquad\times\mathrm{e}^{-\frac{1}{2p}\sum_{\mu=1}^{p}\sum_{i=1}^{N}\sum_{\alpha,\beta=1}^{n}\hat{y}_i^\alpha\hat{y}_i^\beta\theta_\mu^\alpha\theta_\mu^\beta+\mathcal{O}(p^0)} \\
&= -\alpha\log(2\pi)+\lim_{p\to\infty}\lim_{n\to 0}\frac{1}{pn}\log\int\prod_{\alpha=1}^{n}\left(\mathrm{d}z^\alpha\mathrm{d}\boldsymbol{\theta}^\alpha\mathrm{e}^{\mathrm{i}pz^\alpha-\mathrm{i}z^\alpha(\boldsymbol{\theta}^\alpha)^2}\right) \\
&\qquad\times\prod_{i=1}^{N}\int\prod_{\alpha=1}^{n}\left[\mathrm{d}y_i^\alpha\mathrm{d}\hat{y}_i^\alpha\mathrm{e}^{\mathrm{i}\sum_\alpha\hat{y}_i^\alpha y_i^\alpha}\theta[y_i^\alpha-\kappa]\right]\mathrm{e}^{-\frac{1}{2}\sum_{\alpha,\beta}\hat{y}_i^\alpha\hat{y}_i^\beta[\frac{1}{p}\sum_{\mu=1}^{p}\theta_\mu^\alpha\theta_\mu^\beta]}
\end{aligned}
$$

- with $\mathbf{y} = (y_1, \ldots, y_n)$, $\hat{\mathbf{y}} = (\hat{y}_1, \ldots, \hat{y}_n)$, $\mathbf{z} = (z_1, \ldots, z_n)$:

$$\overline{F} = -\alpha \log(2\pi) + \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \int \mathrm{d}\mathbf{z} \Big(\prod_{\alpha=1}^{n} \mathrm{d}\boldsymbol{\theta}^\alpha \mathrm{e}^{\mathrm{i}pz^\alpha - \mathrm{i}z^\alpha(\boldsymbol{\theta}^\alpha)^2} \Big) \times \Big\{ \int \mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\, \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}} \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\, \mathrm{e}^{-\frac{1}{2}\sum_{\alpha,\beta} \hat{y}^\alpha \hat{y}^\beta [\frac{1}{p}\sum_{\mu=1}^{p} \theta_\mu^\alpha \theta_\mu^\beta]} \Big\}^N$$

- insert

$$1 = \int \mathrm{d}q_{\alpha\beta}\, \delta\Big[q_{\alpha\beta} - \frac{1}{p}\sum_{\mu=1}^{p} \theta_\mu^\alpha \theta_\mu^\beta\Big] = \int \frac{\mathrm{d}q_{\alpha\beta}\mathrm{d}\hat{q}_{\alpha\beta}}{2\pi/p}\, \mathrm{e}^{\mathrm{i}p\hat{q}_{\alpha\beta}\left[q_{\alpha\beta} - \frac{1}{p}\sum_{\mu=1}^{p} \theta_\mu^\alpha \theta_\mu^\beta\right]}$$

$$\overline{F} = -\alpha \log(2\pi) + \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \int \mathrm{d}\mathbf{z}\mathrm{d}\mathbf{q}\mathrm{d}\hat{\mathbf{q}}\, \mathrm{e}^{\mathrm{i}p\sum_{\alpha\beta=1}^{n} \hat{q}_{\alpha\beta} q_{\alpha\beta} + \mathrm{i}p\sum_{\alpha=1}^{n} z_\alpha}$$
$$\times \Big\{ \int \mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\, \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}} \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\, \mathrm{e}^{-\frac{1}{2}\hat{\mathbf{y}}\cdot\mathbf{q}\hat{\mathbf{y}}} \Big\}^N$$
$$\times \int \prod_{\alpha=1}^{n} \Big(\mathrm{d}\boldsymbol{\theta}^\alpha \mathrm{e}^{-\mathrm{i}z^\alpha(\boldsymbol{\theta}^\alpha)^2} \Big) \mathrm{e}^{-\mathrm{i}\sum_{\mu=1}^{p} \sum_{\alpha\beta} \hat{q}_{\alpha\beta} \theta_\mu^\alpha \theta_\mu^\beta}$$

- with $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_n)$:
  $(N = \alpha p)$

$$\overline{F} = -\alpha \log(2\pi) + \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \int \mathrm{d}\mathbf{z}\mathrm{d}\mathbf{q}\mathrm{d}\hat{\mathbf{q}}\; \mathrm{e}^{\mathrm{i}p\sum_{\alpha\beta=1}^{n} \hat{q}_{\alpha\beta} q_{\alpha\beta} + \mathrm{i}p\sum_{\alpha=1}^{n} z_\alpha}$$

$$\times \Big\{ \int \mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\; \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}} \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\; \mathrm{e}^{-\frac{1}{2}\hat{\mathbf{y}}\cdot\mathbf{q}\hat{\mathbf{y}}} \Big\}^{\alpha p} \Big\{ \int \mathrm{d}\boldsymbol{\theta}\; \mathrm{e}^{-\mathrm{i}\sum_{\alpha=1}^{n} z^\alpha \theta_\alpha^2 - \mathrm{i}\boldsymbol{\theta}\cdot\hat{\mathbf{q}}\boldsymbol{\theta}} \Big\}^{p}$$

$$= \lim_{p\to\infty} \lim_{n\to 0} \frac{1}{pn} \log \int \mathrm{d}\mathbf{z}\mathrm{d}\mathbf{q}\mathrm{d}\hat{\mathbf{q}}\; \mathrm{e}^{p\Psi(\mathbf{z},\mathbf{q},\hat{\mathbf{q}})}$$

$$\Psi(\ldots) = \mathrm{i}\sum_{\alpha\beta=1}^{n} \hat{q}_{\alpha\beta} q_{\alpha\beta} + \mathrm{i}\sum_{\alpha=1}^{n} z_\alpha + \log \int \mathrm{d}\boldsymbol{\theta}\; \mathrm{e}^{-\mathrm{i}\sum_{\alpha=1}^{n} z^\alpha \theta_\alpha^2 - \mathrm{i}\boldsymbol{\theta}\cdot\hat{\mathbf{q}}\boldsymbol{\theta}}$$

$$+\alpha \log \int \mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\; \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}} \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\; \mathrm{e}^{-\frac{1}{2}\hat{\mathbf{y}}\cdot\mathbf{q}\hat{\mathbf{y}}} - \alpha n \log(2\pi)$$

- assume limits $n \to 0$ and $p \to \infty$ commute,
  steepest descent integration

$$\overline{F} = \lim_{n\to 0} \frac{1}{n} \mathrm{extr}_{\mathbf{z},\mathbf{q},\hat{\mathbf{q}}} \Psi(\mathbf{z}, \mathbf{q}, \hat{\mathbf{q}})$$

Order parameter equations

$$\Psi(\mathbf{z},\mathbf{q},\hat{\mathbf{q}}) = \mathrm{i}\sum_{\alpha\beta=1}^{n}\hat{q}_{\alpha\beta}q_{\alpha\beta} + \mathrm{i}\sum_{\alpha=1}^{n} z_\alpha + \log\int \mathrm{d}\boldsymbol{\theta}\ \mathrm{e}^{-\mathrm{i}\sum_{\alpha=1}^{n} z^\alpha \theta_\alpha^2 - \mathrm{i}\boldsymbol{\theta}\cdot\hat{\mathbf{q}}\boldsymbol{\theta}} + \alpha\log\int \mathrm{d}\mathbf{y}\mathrm{d}\hat{\mathbf{y}}\ \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}}\prod_{\alpha=1}^{n}\theta[y^\alpha-\kappa]\ \mathrm{e}^{-\frac{1}{2}\hat{\mathbf{y}}\cdot\mathbf{q}\hat{\mathbf{y}}} - \alpha n\log(2\pi)$$

- transform $\hat{q}_{\alpha\beta} = -\frac{1}{2}\mathrm{i}k_{\alpha\beta} - z_\alpha\delta_{\alpha\beta}$, integrate over $\hat{\mathbf{y}}$

$$\begin{aligned}\Psi(\mathbf{z},\mathbf{q},\mathbf{k}) &= \frac{1}{2}\sum_{\alpha\beta=1}^{n}k_{\alpha\beta}q_{\alpha\beta} + \mathrm{i}\sum_{\alpha=1}^{n} z_\alpha(1-q_{\alpha\alpha}) + \log\int \mathrm{d}\boldsymbol{\theta}\ \mathrm{e}^{-\frac{1}{2}\boldsymbol{\theta}\cdot\mathbf{k}\boldsymbol{\theta}} \\ &\quad + \alpha\log\int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n}\theta[y^\alpha-\kappa]\int \mathrm{d}\hat{\mathbf{y}}\ \mathrm{e}^{\mathrm{i}\hat{\mathbf{y}}\cdot\mathbf{y}-\frac{1}{2}\hat{\mathbf{y}}\cdot\mathbf{q}\hat{\mathbf{y}}} - \alpha n\log(2\pi) \\ &= \frac{1}{2}\sum_{\alpha\beta=1}^{n}k_{\alpha\beta}q_{\alpha\beta} + \mathrm{i}\sum_{\alpha=1}^{n} z_\alpha(1-q_{\alpha\alpha}) + \log\frac{(2\pi)^{n/2}}{\sqrt{\mathrm{Det}\mathbf{k}}} \\ &\quad + \alpha\log\int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n}\theta[y^\alpha-\kappa]\ \frac{(2\pi)^{n/2}}{\sqrt{\mathrm{Det}\mathbf{q}}}\mathrm{e}^{-\frac{1}{2}\mathbf{y}\cdot\mathbf{q}^{-1}\mathbf{y}} - \alpha n\log(2\pi)\end{aligned}$$

- re-organise

$$\Psi(\mathbf{z},\mathbf{q},\mathbf{k}) = \frac{1}{2}\sum_{\alpha\beta=1}^{n} k_{\alpha\beta}q_{\alpha\beta} + \mathrm{i}\sum_{\alpha=1}^{n} z_\alpha(1-q_{\alpha\alpha}) - \frac{1}{2}\log \mathrm{Det}\,\mathbf{k} - \frac{1}{2}\alpha \log \mathrm{Det}\,\mathbf{q}$$
$$+\ \alpha \log \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{-\frac{1}{2}\mathbf{y}\cdot\mathbf{q}^{-1}\mathbf{y}} + \frac{1}{2}n(1-\alpha)\log(2\pi)$$

- extremise with respect to **z**,
  $\partial\Psi/\partial z_\alpha = 0 :\quad q_{\alpha\alpha} = 1$ for all α

$$\Psi(\mathbf{q},\mathbf{k}) = \frac{1}{2}n(1-\alpha)\log(2\pi) + \frac{1}{2}\sum_{\alpha\beta=1}^{n} k_{\alpha\beta}q_{\alpha\beta} - \frac{1}{2}\log \mathrm{Det}\,\mathbf{k} - \frac{1}{2}\alpha \log \mathrm{Det}\,\mathbf{q}$$
$$+\ \alpha \log \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{-\frac{1}{2}\mathbf{y}\cdot\mathbf{q}^{-1}\mathbf{y}}$$

next: ergodicity assumption,
replica-symmetric form for **q** and **k** ...

Replica symmetric solution

$$\Psi(\mathbf{q},\mathbf{k}) = \frac{1}{2}n(1-\alpha)\log(2\pi) + \frac{1}{2}\sum_{\alpha\beta=1}^{n} k_{\alpha\beta}q_{\alpha\beta} - \frac{1}{2}\log \mathrm{Det}\,\mathbf{k} - \frac{1}{2}\alpha\log\mathrm{Det}\,\mathbf{q}$$
$$+\,\alpha\log\int \mathrm{d}\mathbf{y}\,\prod_{\alpha=1}^{n}\theta[y^{\alpha}-\kappa]\mathrm{e}^{-\frac{1}{2}\mathbf{y}\cdot\mathbf{q}^{-1}\mathbf{y}}$$

RS saddle-points

$$q_{\alpha\beta} = \delta_{\alpha\beta} + (1-\delta_{\alpha\beta})q, \qquad k_{\alpha\beta} = K\delta_{\alpha\beta} + (1-\delta_{\alpha\beta})k$$

- eigenvalues:

$$\mathbf{x} = (1,\ldots,1): \qquad (\mathbf{kx})_{\alpha} = \sum_{\beta=1}^{n}[k+(K-k)\delta_{\alpha\beta}]x_{\beta} = nk+K-k$$
$$\textit{eigenval}: \quad \lambda = nk+K-k$$

$$\sum_{\alpha=1}^{n} x_{\alpha} = 0: \qquad (\mathbf{kx})_{\alpha} = \sum_{\beta=1}^{n}[k+(K-k)\delta_{\alpha\beta}]x_{\beta} = (K-k)x_{\alpha}$$
$$\textit{eigenval}: \quad \lambda = K-k \quad (n-1\ \textit{fold})$$

hence

$$\mathrm{Det}\,\mathbf{k} = (nk+K-k)(K-k)^{n-1}, \qquad \mathrm{Det}\,\mathbf{q} = (nq+1-q)(1-q)^{n-1}$$

- invert $\mathbf{q}$, try $(\mathbf{q}^{-1})_{\alpha\beta} = r + (R-r)\delta_{\alpha\beta}$

$$\begin{aligned}\delta_{\alpha\beta} &= (\mathbf{q}\mathbf{q}^{-1})_{\alpha\beta} = \sum_\gamma (q+(1-q)\delta_{\alpha\gamma})(r+(R-r)\delta_{\gamma\beta}) \\ &= nqr + q(R-r) + r(1-q) + (R-r)(1-q)\delta_{\alpha\beta}\end{aligned}$$

so
$$nqr + q(R-r) + r(1-q) = 0, \qquad (R-r)(1-q) = 1$$
$$R = r + \frac{1}{1-q}, \qquad r = -\frac{q}{(1-q)(1-q+nq)}$$

- hence, using $\exp[\frac{1}{2}x^2] = \int \mathrm{D}z\ \mathrm{e}^{xz}$

$$\begin{aligned}\log \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{-\frac{1}{2}\mathbf{y}\cdot\mathbf{q}^{-1}\mathbf{y}} &= \log \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{-\frac{1}{2}\sum_{\alpha\beta} y_\alpha [r+(R-r)\delta_{\alpha\beta}] y_\beta} \\ &= \log \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{-\frac{1}{2}r[\sum_\alpha y_\alpha]^2 - \frac{1}{2}(R-r)\sum_\alpha y_\alpha^2} \\ &= \log \int \mathrm{D}z \int \mathrm{d}\mathbf{y}\ \prod_{\alpha=1}^{n} \theta[y^\alpha - \kappa]\mathrm{e}^{z\sqrt{-r}\sum_\alpha y_\alpha - \frac{1}{2}(R-r)\sum_\alpha y_\alpha^2} \\ &= \log \int \mathrm{D}z \Big[\int_\kappa^\infty \mathrm{d}y\ \mathrm{e}^{z\sqrt{-r}y - \frac{1}{2}(R-r)y^2}\Big]^n\end{aligned}$$

put everything together ...

$$\begin{aligned}\frac{1}{n}\Psi(\mathbf{q},\mathbf{k}) &= \frac{1}{2}(1-\alpha)\log(2\pi)+\frac{1}{2}K+\frac{1}{2}(n-1)qk-\frac{1}{2n}\log[(nk+K-k)(K-k)^{n-1}]\\ &\quad -\frac{\alpha}{2n}\log[(nq+1-q)(1-q)^{n-1}]+\frac{\alpha}{n}\log\int \mathrm{D}z\Big[\int_\kappa^\infty \mathrm{d}y\ \mathrm{e}^{z\sqrt{-r}y-\frac{1}{2}(R-r)y^2}\Big]^n\\ &= \frac{1}{2}(1-\alpha)\log(2\pi)+\frac{1}{2}(K-qk)-\frac{1}{2n}\log(1+\frac{nk}{K-k})-\frac{1}{2}\log(K-k)\\ &\quad -\frac{\alpha}{2n}\log(1+\frac{nq}{1-q})-\frac{\alpha}{2}\log(1-q)+\mathcal{O}(n)\\ &\quad +\frac{\alpha}{n}\log\int \mathrm{D}z\Big[1+n\log\int_\kappa^\infty \mathrm{d}y\ \mathrm{e}^{zy\sqrt{q}/(1-q)-\frac{1}{2(1-q)}y^2}+\mathcal{O}(n^2)\Big]\end{aligned}$$

take $n\to 0$:

$$\begin{aligned}2\overline{F} &= (1-\alpha)\log(2\pi)+\mathrm{extr}_{K,k,q}\Big\{K-qk-\frac{k}{K-k}-\log(K-k)\\ &\quad -\frac{\alpha q}{1-q}-\alpha\log(1-q)+2\alpha\int \mathrm{D}z\ \log\int_\kappa^\infty \mathrm{d}y\ \mathrm{e}^{zy\sqrt{q}/(1-q)-\frac{1}{2(1-q)}y^2}\Big\}\end{aligned}$$

$$2\overline{F} = (1-\alpha)\log(2\pi) + \mathrm{extr}_{K,k,q}\Big\{ K - qk - \frac{k}{K-k} - \log(K-k) - \frac{\alpha q}{1-q} - \alpha\log(1-q) + 2\alpha \int \mathrm{D}z \, \log \int_{\kappa}^{\infty} \mathrm{d}y \, \mathrm{e}^{zy\sqrt{q}/(1-q) - \frac{1}{2(1-q)}y^2} \Big\}$$

- extremise over K and k

$$\begin{cases} \frac{\partial}{\partial K} = 0: & 1 + \frac{k}{(K-k)^2} - \frac{1}{K-k} = 0 \\ \frac{\partial}{\partial k} = 0: & -q - \frac{1}{K-k} - \frac{k}{(K-k)^2} + \frac{1}{K-k} = 0 \end{cases} \Rightarrow \quad K = \frac{1-2q}{(1-q)^2}, \quad k = -\frac{q}{(1-q)^2}$$

$$2\overline{F} = (1-\alpha)\log(2\pi) + \mathrm{extr}_{q}\Big\{ \frac{1}{1-q} - \frac{\alpha q}{1-q} + (1-\alpha)\log(1-q) + 2\alpha \int \mathrm{D}z \, \log \int_{\kappa}^{\infty} \mathrm{d}y \, \mathrm{e}^{zy\sqrt{q}/(1-q) - \frac{1}{2(1-q)}y^2} \Big\}$$

- in terms of error function, $\mathrm{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{d}t \, \mathrm{e}^{-t^2}$:

$$\int_\kappa^\infty \mathrm{d}y \, \mathrm{e}^{zy\sqrt{q}/(1-q) - \frac{1}{2(1-q)} y^2} = \mathrm{e}^{\frac{qz^2}{2(1-q)}} \int_\kappa^\infty \mathrm{d}y \, \mathrm{e}^{-\frac{[y - z\sqrt{q}]^2}{2(1-q)}}$$

$$= \sqrt{2(1-q)} \, \mathrm{e}^{\frac{qz^2}{2(1-q)}} \frac{\sqrt{\pi}}{2} \left\{ 1 - \mathrm{Erf}\left[\frac{K - z\sqrt{q}}{\sqrt{2(1-q)}} \right] \right\}$$

$$2\overline{F} = \log \pi + (1-2\alpha) \log 2 + \mathrm{extr}_q \left\{ \frac{1}{1-q} + \log(1-q) + 2\alpha \int \mathrm{D}z \, \log \left[1 - \mathrm{Erf}\left(\frac{K - z\sqrt{q}}{\sqrt{2(1-q)}} \right) \right] \right\}$$

- extremisation with respect to q

short-hand $u(z,q) = (\kappa - z\sqrt{q})/\sqrt{2(1-q)}$,
use $\mathrm{Erf}'(x) = \frac{2}{\sqrt{\pi}} \exp[-x^2]$

$$\frac{\mathrm{d}}{\mathrm{d}q} = 0 : \qquad \frac{1}{(1-q)^2} - \frac{1}{1-q} - 2\alpha \int \mathrm{D}z \left(\frac{\partial u}{\partial q} \right) \frac{\mathrm{Erf}' \; u(z,q)}{1 - \mathrm{Erf} \; u(z,q)} = 0$$

- work out further

$$\frac{q}{(1-q)^2} = \frac{4\alpha}{\sqrt{\pi}} \int \mathrm{D}z \left(\frac{\partial u}{\partial q}\right) \frac{\mathrm{e}^{-u^2(z,q)}}{1-\mathrm{Erf}\; u(z,q)}$$

$$\frac{\partial u}{\partial q} = \frac{1}{\sqrt{2}} \frac{\partial}{\partial q} \frac{\kappa - z\sqrt{q}}{(1-q)^{1/2}} = \ldots = \frac{\kappa\sqrt{q} - z}{2\sqrt{2q}(1-q)^{3/2}}$$

insert:

$$q\sqrt{q} = \alpha\sqrt{\frac{2}{\pi}}\sqrt{1-q} \int \mathrm{D}z\; \frac{\mathrm{e}^{-u^2(z,q)}(\kappa\sqrt{q}-z)}{1-\mathrm{Erf}\; u(z,q)}$$

$$2\overline{F} = \log\pi + (1-2\alpha)\log 2 + \frac{1}{1-q} + \log(1-q) + 2\alpha \int \mathrm{D}z\; \log\left[1-\mathrm{Erf}\; u(z,q)\right]$$

$$q\sqrt{q} = \alpha\sqrt{\frac{2}{\pi}}\sqrt{1-q} \int \mathrm{D}z\; \frac{\mathrm{e}^{-u^2(z,q)}(\kappa\sqrt{q}-z)}{1-\mathrm{Erf}\; u(z,q)}, \qquad u(z,q) = \frac{\kappa - z\sqrt{q}}{\sqrt{2(1-q)}}$$

remember:

$\overline{F}$ =finite: random data linearly separable with margin κ

$\overline{F} = -\infty$: random data not linearly separable with margin κ

- $\alpha = 0$ (so $1 \ll N \ll p$):

  $q = 0, \quad 2\overline{F} = \log\pi + \log 2 + 1$ *random data linearly separable* (*overfitting*)

- $\alpha > 0$ (so $1 \ll N \sim p$):

  transition point: value of α where $q \to 1$

$$1 = \alpha_c(\kappa)\sqrt{\frac{2}{\pi}}\int \mathrm{D}z \lim_{q\to 1}\sqrt{1-q}\,\frac{\mathrm{e}^{-[\frac{K-z}{\sqrt{2(1-q)}}]^2}(\kappa - z)}{1-\mathrm{Erf}\left[\frac{\kappa-z}{\sqrt{2(1-q)}}\right]}$$

$$\alpha_c(\kappa) = \left[\frac{1}{\sqrt{\pi}}\int \mathrm{D}z\,(\kappa+z)\lim_{\gamma\to\infty}\frac{1}{\gamma}\,\frac{\mathrm{e}^{-\gamma^2(\kappa+z)^2}}{1-\mathrm{Erf}[\gamma(\kappa+z)]}\right]^{-1}$$

- remaining limit:

$$\lim_{\gamma\to\infty} \frac{1}{\gamma} \frac{\mathrm{e}^{-\gamma^2 Q^2}}{1-\mathrm{Erf}[\gamma Q]} = Q\sqrt{\pi}\,\theta(Q)$$

proof:

$Q<0$: $\quad \mathrm{Erf}[\gamma Q] \to -1$ *so* $\lim_{\gamma\to\infty} \frac{1}{\gamma} \frac{\mathrm{e}^{-\gamma^2 Q^2}}{1-\mathrm{Erf}[\gamma Q]} = 0$

$Q>0$: $\quad \mathrm{Erf}[\gamma Q] = 1 - \frac{1}{\gamma Q\sqrt{\pi}} \mathrm{e}^{-\gamma^2 Q^2}\Big(1+\mathcal{O}(\frac{1}{\gamma^2 Q^2})\Big)$

$$\lim_{\gamma\to\infty} \frac{1}{\gamma} \frac{\mathrm{e}^{-\gamma^2 Q^2}}{1-\mathrm{Erf}[\gamma Q]} = \lim_{\gamma\to\infty} \frac{1}{\gamma} \frac{\mathrm{e}^{-\gamma^2 Q^2}}{\frac{1}{\gamma Q\sqrt{\pi}}\mathrm{e}^{-\gamma^2 Q^2}\Big(1+\mathcal{O}(\frac{1}{\gamma^2 Q^2})\Big)} = Q\sqrt{\pi}$$

Final result

$$\alpha_c(\kappa) = \left[\int_{-\kappa}^{\infty} \mathrm{D}z\, (\kappa+z)^2\right]^{-1} \qquad \alpha_c(0) = \left[\int_0^{\infty} \mathrm{D}z\, z^2\right]^{-1} = [\tfrac{1}{2}]^{-1} = 2$$

p covariates, *N* patients
binary outcomes,
p and *N* large

<u>random</u> data
(i.e. pure binary noise)
is *perfectly* separable if
$N/p < \alpha_c(\kappa)$

machine learning algorithms
will find pararameters $\theta_1 \ldots \theta_p$
such that $t_i = sgn[\sum_{\mu=1}^{p} \theta_\mu x_\mu^i]$
for all $i = 1 \ldots N$