Discrete symmetries in statistical image analysis and beyond

Alexey Koloydenko

Department of Mathematics
Royal Holloway University of London

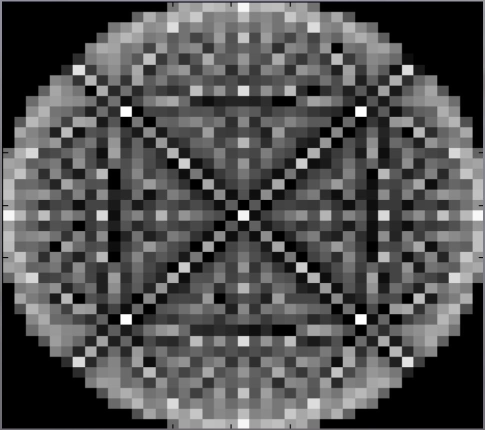
21st January 2009

A natural image, collection of van Hateren (J. H. van Hateren, A. van der Schaaf 1998)

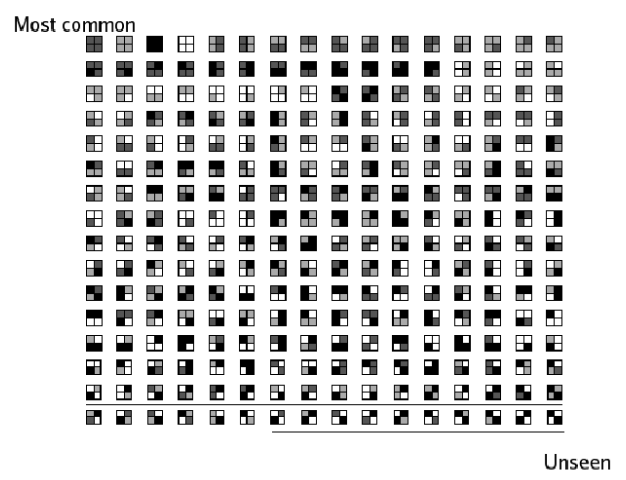
A joint histogram of contrast-normalized differences

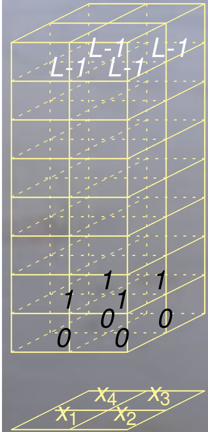
$$(x_4 + x_3 - x_1 - x_2, x_4 + x_1 - x_3 - x_2)\,/\,c(\mathbf{x})$$

within microimage $\mathbf{x} = \begin{matrix} x_4 & x_3 \\ x_1 & x_2 \end{matrix}$

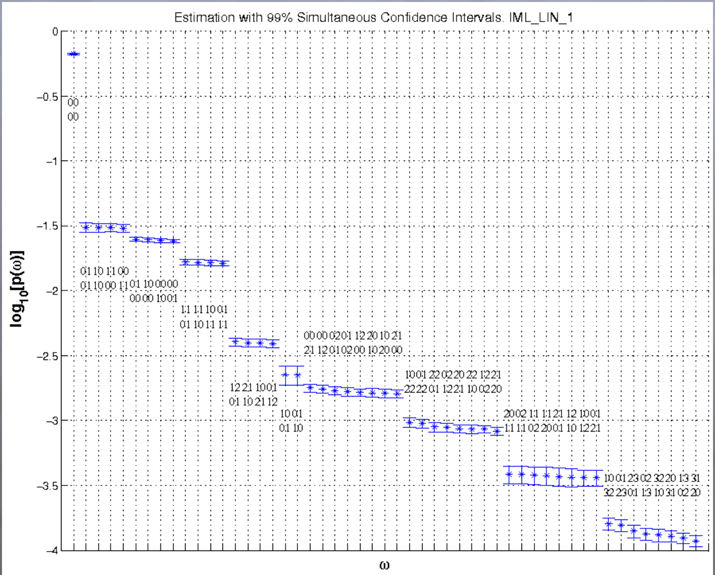


Modes: $\left\{ \begin{matrix} 0 & 0 \\ 0 & 0 \end{matrix}, \begin{matrix} 1 & 0 \\ 0 & 1 \end{matrix}, \begin{matrix} 0 & 1 \\ 1 & 0 \end{matrix} \right\}$; $\begin{matrix} 1 & 0 \\ 0 & 0 \end{matrix}, \begin{matrix} 0 & 1 \\ 0 & 0 \end{matrix}, \begin{matrix} 0 & 0 \\ 0 & 1 \end{matrix}, \begin{matrix} 0 & 0 \\ 1 & 0 \end{matrix}$; $\begin{matrix} 1 & 0 \\ 1 & 0 \end{matrix}, \begin{matrix} 0 & 1 \\ 0 & 1 \end{matrix}, \begin{matrix} 1 & 1 \\ 0 & 0 \end{matrix}, \begin{matrix} 0 & 0 \\ 1 & 1 \end{matrix}$





Micropatterns $\mathbf{x} \in \Omega = \mathrm{Mat}_{2\times 2}([L])$, $L = 4$, ranked by frequency. Symmetry group of square cuboid: $G = D_{4h} \cong D_8 \times C_2$ (Geman, Koloydenko 1998; Koloydenko, Geman 2006)



Estimation with 99% Simultaneous Confidence Intervals. IML_LIN_1
0
−0.5
−1
−1.5
−2
−2.5
−3
−3.5
−4
log10[p(ω)]
ω
00
00
01 10 11 00
01 10 00 11
01 10 00 00
00 00 10 01
11 11 10 01
01 10 11 11
00 00 02 01 12 20 10 21
21 12 01 02 00 10 20 00
12 21 10 01
01 10 21 12
10 01
01 10
10 01 22 02 20 22 12 21
22 22 01 12 21 10 02 20
20 02 11 11 21 12 10 01
11 11 02 20 01 10 12 21
10 01 23 02 32 20 13 31
32 23 01 13 10 31 02 20

Full Group G of Microimage Symmetries

$$p(\mathbf{x}) = p(g\mathbf{x}), \quad \forall \mathbf{x} \in \Omega, \ \forall g \in G$$

$$r = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad s = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad i = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Presentation of
$G = \langle r, s, i | r^4 = s^2 = i^2 = 1, si = is, ri = ir, rs = sr^3 \rangle$.

Note: Ω is translated down to be centered around $\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}$ in order for i to act as $-1\times$

Full Group G of Microimage Symmetries

$$p(\boldsymbol{x}) = p(g\boldsymbol{x}), \quad \forall \boldsymbol{x} \in \Omega,\ \forall g \in G$$

$$r = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad s = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \quad i = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

Presentation of
$G = \langle r, s, i | r^4 = s^2 = i^2 = 1, si = is, ri = ir, rs = sr^3 \rangle$.

Note: Ω is translated down to be centered around $\begin{smallmatrix} 0 & 0 \\ 0 & 0 \end{smallmatrix}$ in order for i to act as $-1\times$

- A key to modelling of G-invariant p (for any finite group G) is *algebra (ring)* of G-invariant polynomials. In our case, $\mathbb{R}[x_1, x_2, x_3, x_4]^G$, where
- G-action on $\mathbb{R}[x_1, x_2, x_3, x_4]$ is defined as follows:

$$rx_1 = x_2; \quad rx_2 = x_3; \quad rx_3 = x_4; \quad rx_4 = x_1;$$
$$sx_1 = x_1; \quad sx_2 = x_4; \quad sx_3 = x_3; \quad sx_4 = x_2;$$
$$ix_k = -x_k, \; k = 1, 2, 3, 4$$

- *Theory of polynomial invariants of finite groups* (D. Hilbert 1890, E. Noether 1916):
- $\mathbb{R}[x_1, x_2, \ldots, x_m]^G$ has a finite set of generators $f_1(x_1, x_2, \ldots, x_m), f_2(x_1, x_2, \ldots, x_m), \ldots, f_N(x_1, x_2, \ldots, x_m)$

- A key to modelling of G-invariant p (for any finite group G) is *algebra (ring)* of G-invariant polynomials. In our case, $\mathbb{R}[x_1, x_2, x_3, x_4]^G$, where
- G-action on $\mathbb{R}[x_1, x_2, x_3, x_4]$ is defined as follows:

$$rx_1 = x_2; \quad rx_2 = x_3; \quad rx_3 = x_4; \quad rx_4 = x_1;$$
$$sx_1 = x_1; \quad sx_2 = x_4; \quad sx_3 = x_3; \quad sx_4 = x_2;$$
$$ix_k = -x_k, \; k = 1, 2, 3, 4$$

- *Theory of polynomial invariants of finite groups* (D. Hilbert 1890, E. Noether 1916):
- $\mathbb{R}[x_1, x_2, \ldots, x_m]^G$ has a finite set of generators $f_1(x_1, x_2, \ldots, x_m), f_2(x_1, x_2, \ldots, x_m), \ldots, f_N(x_1, x_2, \ldots, x_m)$

- A key to modelling of G-invariant p (for any finite group G) is *algebra (ring)* of G-invariant polynomials. In our case, $\mathbb{R}[x_1, x_2, x_3, x_4]^G$, where
- G-action on $\mathbb{R}[x_1, x_2, x_3, x_4]$ is defined as follows:

$$rx_1 = x_2; \quad rx_2 = x_3; \quad rx_3 = x_4; \quad rx_4 = x_1;$$
$$sx_1 = x_1; \quad sx_2 = x_4; \quad sx_3 = x_3; \quad sx_4 = x_2;$$
$$ix_k = -x_k, \; k = 1, 2, 3, 4$$

- *Theory of polynomial invariants of finite groups* (D. Hilbert 1890, E. Noether 1916):
- $\mathbb{R}[x_1, x_2, \ldots, x_m]^G$ has a finite set of generators $f_1(x_1, x_2, \ldots, x_m), f_2(x_1, x_2, \ldots, x_m), \ldots, f_N(x_1, x_2, \ldots, x_m)$

- A key to modelling of G-invariant p (for any finite group G) is *algebra (ring)* of G-invariant polynomials. In our case, $\mathbb{R}[x_1, x_2, x_3, x_4]^G$, where
- G-action on $\mathbb{R}[x_1, x_2, x_3, x_4]$ is defined as follows:

$$rx_1 = x_2; \quad rx_2 = x_3; \quad rx_3 = x_4; \quad rx_4 = x_1;$$
$$sx_1 = x_1; \quad sx_2 = x_4; \quad sx_3 = x_3; \quad sx_4 = x_2;$$
$$ix_k = -x_k, \ k = 1, 2, 3, 4$$

- *Theory of polynomial invariants of finite groups* (D. Hilbert 1890, E. Noether 1916):
- $\mathbb{R}[x_1, x_2, \ldots, x_m]^G$ has a finite set of generators $f_1(x_1, x_2, \ldots, x_m), f_2(x_1, x_2, \ldots, x_m), \ldots, f_N(x_1, x_2, \ldots, x_m)$

Theorem

$$\begin{aligned} f_1(x) &= (x_1 + x_3)(x_2 + x_4), \\ f_2(x) &= x_1 x_3 + x_2 x_4, \\ f_3(x) &= x_1^2 + x_2^2 + x_3^2 + x_4^2, \\ f_4(x) &= x_1 x_2 x_3 x_4, \\ f_5(x) &= (x_1^2 + x_3^2)(x_2^2 + x_4^2). \end{aligned}$$

Also,

$$\mathbb{R}[x_1, x_2, x_3, x_4]^G \overset{(f_1,\dots,f_5)}{\cong} \mathbb{R}[w_1, w_2, w_3, w_4, w_5]/J_F, \text{ where}$$

$$J_F = \{h \in \mathbb{R}[w_1, w_2, w_3, w_4, w_5] : h(f_1, f_2, f_3, f_4, f_5) = 0 \in \mathbb{R}[x_1, x_2, x_3, x_4]\} = \langle q \rangle, \text{ and}$$

$$q(w_1, w_2, w_3, w_4, w_5) = 4w_1^2 w_3 + 8w_1 w_2 w_5 + 2w_1 w_3 w_5 - 2w_1 w_4^2 w_5 + \\ +16w_2^2 - 8w_2 w_3 - 8w_2 w_4^2 + 4w_2 w_5^2 + w_3^2 - 2w_3 w_4^2 + w_4^4.$$

Theorem

$$\begin{aligned} f_1(x) &= (x_1+x_3)(x_2+x_4),\\ f_2(x) &= x_1x_3+x_2x_4,\\ f_3(x) &= x_1^2+x_2^2+x_3^2+x_4^2,\\ f_4(x) &= x_1x_2x_3x_4,\\ f_5(x) &= (x_1^2+x_3^2)(x_2^2+x_4^2). \end{aligned}$$

Also,

$$\mathbb{R}[x_1, x_2, x_3, x_4]^G \overset{(f_1,\ldots,f_5)}{\cong} \mathbb{R}[w_1, w_2, w_3, w_4, w_5]/J_F, \text{ where}$$

$$J_F = \{h \in \mathbb{R}[w_1, w_2, w_3, w_4, w_5] : h(f_1, f_2, f_3, f_4, f_5) = 0 \in \mathbb{R}[x_1, x_2, x_3, x_4]\} = \langle q \rangle, \text{ and}$$

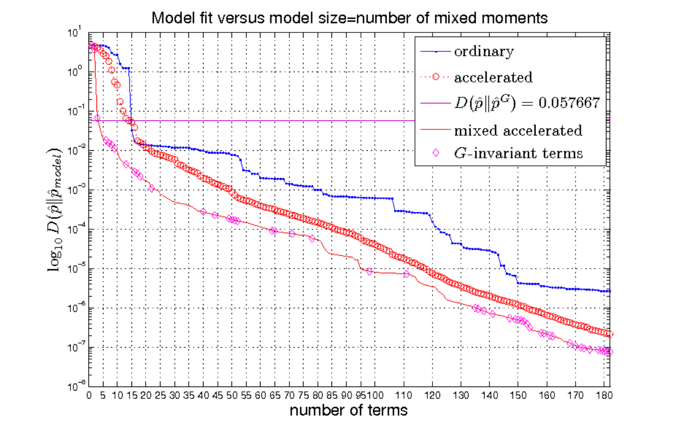
$$q(w_1, w_2, w_3, w_4, w_5) = 4w_1^2w_3 + 8w_1w_2w_5 + 2w_1w_3w_5 - 2w_1w_4^2w_5 + \\ +16w_2^2 - 8w_2w_3 - 8w_2w_4^2 + 4w_2w_5^2 + w_3^2 - 2w_3w_4^2 + w_4^4.$$

General measure: ordinary mixed moments		G-invariant measure: *mixed G-invariant moments*
$x_1^{\alpha_1} x_2^{\alpha_2} \dots x_m^{\alpha_m}$	$\leftrightarrow$	$f_1^{\beta_1} f_2^{\beta_2} \dots f_N^{\beta_N}$

Symbolic algebra tools (e.g. Gap, INVAR, Macaulay2, Magma) are readily available to produce generators $f_1, f_2, \dots, f_N$ (and J_F, relations on them).

General measure: ordinary mixed moments		*G*-invariant measure: *mixed G-invariant moments*
$x_1^{\alpha_1} x_2^{\alpha_2} \ldots x_m^{\alpha_m}$	$\leftrightarrow$	$f_1^{\beta_1} f_2^{\beta_2} \ldots f_N^{\beta_N}$

Symbolic algebra tools (e.g. Gap, INVAR, Macaulay2, Magma) are readily available to produce generators $f_1, f_2, \ldots, f_N$ (and J_F, relations on them).



Log-linear models with increasing number of ordinary and *G*-invariant terms are fit to natural microimage data. The maximal *G*-invariant model has 30 free parameters but its four parameter simplification with three nonconstant *G*-invariant terms (f_3, f_1, and f_3^2) nearly achieves the best *G*-invariant fit. These together with *G*-invariant terms ($f_1 f_3, f_4, f_3^3$) are also included by greedy ("accelerated") constructor in the best-fitting ten parameter model.

Implications for statistical modelling

1. Equality constraints $\leftrightarrow$ structure of Ω
2. Ordered levels $[L] = \{1, 2, \ldots, L\} \Rightarrow \Omega \subset \mathbb{R}^m$
3. Structure: $\{\mathcal{O}\} \cong \Omega / G$, G - a subgroup of the full symmetry group of Ω
4. Multiple data sets of common origin (common constraints) but variable resolution (quantisation) $\Rightarrow$ multiple Ω's identified with common solid
5. Statistical models should be compatible (e.g. admit appropriate extensions (P. McCullagh 2002) with progressive refining of Ω)

Implications for statistical modelling

1. Equality constraints $\leftrightarrow$ structure of Ω
2. Ordered levels $[L] = \{1, 2, \ldots, L\} \Rightarrow \Omega \subset \mathbb{R}^m$
3. Structure: $\{\mathcal{O}\} \cong \Omega / G$, G - a subgroup of the full symmetry group of Ω
4. Multiple data sets of common origin (common constraints) but variable resolution (quantisation) $\Rightarrow$ multiple Ω's identified with common solid
5. Statistical models should be compatible (e.g. admit appropriate extensions (P. McCullagh 2002) with progressive refining of Ω)

Implications for statistical modelling

1. Equality constraints $\leftrightarrow$ structure of Ω
2. Ordered levels $[L] = \{1, 2, \ldots, L\} \Rightarrow \Omega \subset \mathbb{R}^m$
3. Structure: $\{\mathcal{O}\} \cong \Omega / G$, G - a subgroup of the full symmetry group of Ω
4. Multiple data sets of common origin (common constraints) but variable resolution (quantisation) $\Rightarrow$ multiple Ω's identified with common solid
5. Statistical models should be compatible (e.g. admit appropriate extensions (P. McCullagh 2002) with progressive refining of Ω)

Implications for statistical modelling

1. Equality constraints $\leftrightarrow$ structure of Ω
2. Ordered levels $[L] = \{1, 2, \ldots, L\} \ \Rightarrow \Omega \subset \mathbb{R}^m$
3. Structure: $\{\mathcal{O}\} \cong \Omega/G$, G - a subgroup of the full symmetry group of Ω
4. Multiple data sets of common origin (common constraints) but variable resolution (quantisation) $\Rightarrow$ multiple Ω's identified with common solid
5. Statistical models should be compatible (e.g. admit appropriate extensions (P. McCullagh 2002) with progressive refining of Ω)

Implications for statistical modelling

1. Equality constraints $\leftrightarrow$ structure of Ω
2. Ordered levels $[L] = \{1, 2, \ldots, L\} \Rightarrow \Omega \subset \mathbb{R}^m$
3. Structure: $\{\mathcal{O}\} \cong \Omega / G$, G - a subgroup of the full symmetry group of Ω
4. Multiple data sets of common origin (common constraints) but variable resolution (quantisation) $\Rightarrow$ multiple Ω's identified with common solid
5. Statistical models should be compatible (e.g. admit appropriate extensions (P. McCullagh 2002) with progressive refining of Ω)

Summary

1. Generalisation of the multivariate *Problem of moments*
2. A new connection with computational algebraic geometry
3. Incremental model construction in the presence of G-invariance
4. Reduction of computations for estimation

Extensions

G-invariant polynomials are "all" patch features invariant to G. Hence, any (possibly non-linear) regression on G-invariant features, or predictors, should be carried out in terms of $f_1, f_2, \ldots, f_N$. Besides modeling the microimage distributions p (possibly conditional on membership of the patch in some class of interest), the response might naturally be probability of membership of the patch in some class of interest.

Summary

1. Generalisation of the multivariate *Problem of moments*
2. A new connection with computational algebraic geometry
3. Incremental model construction in the presence of G-invariance
4. Reduction of computations for estimation

Extensions

G-invariant polynomials are “all” patch features invariant to G. Hence, any (possibly non-linear) regression on G-invariant features, or predictors, should be carried out in terms of $f_1, f_2, \ldots, f_N$. Besides modeling the microimage distributions p (possibly conditional on membership of the patch in some class of interest), the response might naturally be probability of membership of the patch in some class of interest.

Summary

1. Generalisation of the multivariate *Problem of moments*
2. A new connection with computational algebraic geometry
3. Incremental model construction in the presence of G-invariance
4. Reduction of computations for estimation

Extensions

G-invariant polynomials are “all” patch features invariant to G. Hence, any (possibly non-linear) regression on G-invariant features, or predictors, should be carried out in terms of $f_1, f_2, \ldots, f_N$. Besides modeling the microimage distributions p (possibly conditional on membership of the patch in some class of interest), the response might naturally be probability of membership of the patch in some class of interest.

Summary

1. Generalisation of the multivariate *Problem of moments*
2. A new connection with computational algebraic geometry
3. Incremental model construction in the presence of G-invariance
4. Reduction of computations for estimation

Extensions

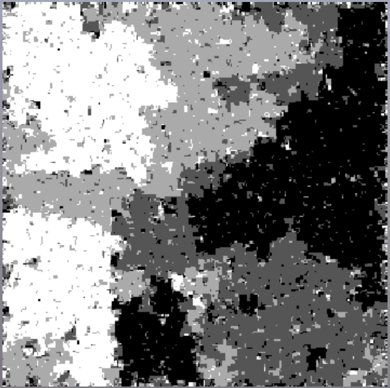
G-invariant polynomials are "all" patch features invariant to G. Hence, any (possibly non-linear) regression on G-invariant features, or predictors, should be carried out in terms of $f_1, f_2, \ldots, f_N$. Besides modeling the microimage distributions p (possibly conditional on membership of the patch in some class of interest), the response might naturally be probability of membership of the patch in some class of interest.

Summary

1. Generalisation of the multivariate *Problem of moments*
2. A new connection with computational algebraic geometry
3. Incremental model construction in the presence of G-invariance
4. Reduction of computations for estimation

Extensions

G-invariant polynomials are "all" patch features invariant to G. Hence, any (possibly non-linear) regression on G-invariant features, or predictors, should be carried out in terms of $f_1, f_2, \ldots, f_N$. Besides modeling the microimage distributions p (possibly conditional on membership of the patch in some class of interest), the response might naturally be probability of membership of the patch in some class of interest.

Left: The natural image with $L = 4$ grey levels. Right: A realization of the Maximum Entropy extension of the G-invariant estimates of the natural 2×2 microimage distribution (courtesy of Prof. L. Younes, the Johns Hopkins University)

Reference

"Symmetric Measures via Moments" *Bernoulli* Volume 14, Number 2, pp. 362–390, 2008